Beiträge zur Untersuchung von Merkmalsfolgen kollektivähnlichen Verhaltens

Inaugural-Dissertation
zur Erlangung der Doktorwürde
einer Hohen Naturwissenschaftlichen Fakultät
der Johann Wolfgang Goethe-Universität
zu Frankfurt a. Main

vorgelegt von
Gerhard Schauffler
aus Stuttgart

Druck von W. Kohlhammer · Stuttgart 1940

Von der Naturwissenschaftlichen Fakultät der
Johann Wolfgang Goethe-Universität zu Frankfurt a. Main
angenommen den 23. September 1939

Berichterstatter: Professor Dr. G. Aumann

Tag der mündlichen Prüfung: 3. Oktober 1939

R. UNIV.
BIBLIOTHEEK
LEIDEN

Einleitung

Ähnliche Idealisierungen der Wirklichkeit, wie sie bei der Aufstellung geometrischer Begriffe vorgenommen werden müssen, sind auch für die Einführung des Wahrscheinlichkeitsbegriffs notwendig; und genau wie die Geometrie mit einem Mindestmaß solcher Idealisierungen auszukommen sucht, erscheint es auch in der Wahrscheinlichkeitsrechnung mit Rücksicht auf die Anwendungen wünschenswert, die erforderlichen Voraussetzungen möglichst weitgehend einzuschränken.

Liegt eine genügend große endliche Anzahl von Ergebnissen statistischer Versuche vor, so besteht eine erste Idealisierung dieser endlichen Reihe darin, daß man sich die Reihe unbegrenzt fortgesetzt denkt. Ordnet man nun jedem Glied der Reihe die bis an die betreffende Stelle gemessene relative Häufigkeit für das Vorkommen eines Merkmals m zu, so beschäftigt sich die Wahrscheinlichkeitsrechnung vorwiegend mit der Folge dieser relativen Häufigkeiten. Diese allgemeine Folge schränkt nun die Wahrscheinlichkeitsrechnung durch Konvergenzforderungen ein: Die Folge der relativen Häufigkeiten und gewisse daraus abgeleitete Folgen sollen konvergent sein.

Nun läßt sich aber bei der Betrachtung statistischer Reihen aus dem vorliegenden endlichen Stück an sich gar nicht auf das Vorhandensein einer Konvergenz der Folge der relativen Häufigkeiten schließen. Es erscheint daher für viele — namentlich für die Anwendungen der Wahrscheinlichkeitsrechnung wichtige — Fälle zweckmäßig, die in der Praxis nicht nachprüfbare Forderung der Existenz des Grenzwerts fallen zu lassen. Dies wird auch im ersten Teil der vorliegenden Arbeit geschehen. Dabei wollen wir jedoch auf eine Bedingung nicht verzichten. Es soll nämlich jede endliche Merkmalkombination in der betrachteten unendlichen Merkmalsfolge unendlich oft auftreten (Definitionseigenschaft der später als X-Folgen bezeichneten E-Folgen); man kann dies als eine Art von „Regellosigkeit" ansehen.

In der vorliegenden Arbeit wird untersucht, welche Gestalt bei Verzicht auf die Konvergenzbedingung, Sätze der Wahrscheinlichkeitstheorie annehmen, insbesondere die beiden Aussagen des Bernoullischen Satzes.

Zunächst werden im Zuge der in dieser Arbeit zugrunde gelegten Auffassung von möglichster Befreiung von Existenzvoraussetzungen über Grenzwerte für die W-Bänder und V-Bereiche allgemeiner Merkmalsfolgen einige Sätze abgeleitet, die als Anfang einer „direkten" Wahrscheinlichkeitstheorie aufgefaßt werden können.

Die Form, in der die erste Aussage des Bernoullischen Gesetzes für sich überdeckende Serien im Sinne der Wahrscheinlichkeitstheorie der Merkmalsfolgen gegeben wird, kann man als die allgemeinste Formulierung dieser Aussage überhaupt ansehen. Bei Beschränkung auf „Bernoullische Folgen" ergibt sich daraus die klassische Form des Gesetzes.

Für die zweite Bernoullische Aussage wird nur in einem Spezialfall eine „direkte" Formulierung gegeben, d. h. also eine Formulierung, bei der in möglichst weitgehendem Maße auf die Voraussetzungen der Existenz von Grenzwerten verzichtet wird. Die Untersuchungen lassen aber deutlich die Schwierigkeiten erkennen, die hier der Behandlung des allgemeinen Falles entgegentreten.

Im zweiten Teil der Arbeit werden dann die sogenannten nachwirkungsfreien Folgen behandelt. Für diese wird ein bedeutungsvoller Invarianzsatz bewiesen, daß nämlich diese Folgen bei Auswahl durch Vorgängerkombination wieder auf nachwirkungsfreie Folgen führen, eine Eigenschaft, die gelten muß, wenn man das „Prinzip vom ausgeschlossenen Spielsystem" fordert.

Auf Grund eines elementaren Satzes der Kombinatorik wird ein Existenzbeweis für die Z-Folgen gegeben, die im wesentlichen mit den nachwirkungsfreien Folgen identisch sind.

Mit Rücksicht auf etwaige Beobachtungsfehler wird schließlich bei allen betrachteten Folgenarten die Unempfindlichkeit der Folge gegenüber mit der Grenzhäufigkeit O — also in Endlichen sehr selten — vorkommenden Streichungen oder Abänderungen untersucht.

I. X - Folgen.

§ 1. E - Folgen.

In den Anwendungsgebieten der Wahrscheinlichkeitsrechnung sollen stets aus einer grossen Anzahl von Versuchen, Beobachtungen, Tatbeständen oder sonst irgend welchen Dingen gewisse zahlenmässige Eigenschaften abgeleitet werden. In der einfachsten Form lautet die Fragestellung, wie oft eine bestimmte Eigenschaft, ein bestimmtes Merkmal vorkommt. Ueberträgt man diese Fragestellung aufdie - als Idealisierung - unbegrenzt fortgesetzt gedachte Reihe, so erhält man die Definition der E-Folge:

Def.I. Unter einer E-Folge wird eine (einfach) unendliche Folge von bestimmten Elementen oder Gliedern verstanden, wenn für jedes Element der E-Folge entscheidbar ist, ob ihm ein gegebenes Merkmal m zukommt oder nicht.

Nach dieser Definition ist es ganz gleichgültig, um was es sich bei den Elementen der Folge handelt; es kommt nur darauf an, ob den Elementen ein bestimmtes Merkmal zukommt oder nicht; man kann deshalb die Elemente jederzeit z.B. durch die Zahlen 1 oder 0 ersetzen, je nachdem ob dem betreffenden Element das Merkmal zukommt oder nicht.

Vielfach interessiert uns nicht nur ein einziges Merkmal m, sondern endlich oder abzählbar unendlich viele Merkmale m_1 , m_2 , m_3 , ... Zwei Merkmale m_1 und m_2 heissen einander ausschliessend, wenn keinem Element der Folge beide Merkmale zukommen können. In der Folge der natürlichen Zahlen sind z.B. die Merkmale " gerade " und " ungerade " einander ausschliessende Merkmale, während dies für die Merkmale " teilbar durch 2 " und " teilbar durch 3 " nicht der Fall ist.

Neben der Frage, ob einem Element der E-Folge ein Merkmal m zukommt oder nicht, ist die Stellennummer des Elements von Bedeutung. Die Anzahl derjenigen unter den n ersten Gliedern der E-Folge E, auf welche das Merkmal m zutrifft, heisst Häufigkeit und wird mit $H_n(E,m)$ bezeichnet. Durch Division mit der Anzahl n erhält man die relative Häufigkeit $h_n (E,m)$.[1)]

Da $0 \leqq H_n(E) \leqq n$, also $0 \leqq h_n(E) \leqq 1$ ist, so ist die Zahlenfolge $h_1(E)$, $h_2(E)$, $h_3(E)$, ... beschränkt. Nach dem Satz von Bolzano - Weierstrass existieren also :

$$\underline{\lim_{n=\infty}} h_n(E) = g \quad ; \quad \overline{\lim_{n=\infty}} h_n(E) = G \quad .$$

---

1) Gelegentlich schreiben wir statt $H_n(E,m)$ und $h_n(E,m)$ kürzer $H_n(E)$ und $h_n(E)$, wenn in Bezug auf das Merkmal m ein Missverständnis ausgeschlossen ist.

Ueber die E-Folgen hat Kamke[2] folgenden Satz bewiesen :

Satz 1. Im Intervall $g \leqq w \leqq G$ liegen die relativen Häufigkeiten überall dicht.

Dieser Satz bildet die Grundlage für den Begriff des W-Bandes :

Def.II. Das Intervall $g \leqq w \leqq G$ heisst Wahrscheinlichkeitsband, abgekürzt W-Band und wird mit B (E,m) bezeichnet.

Aus dieser Definition folgt nach Satz 1, dass jede E-Folge ein W-Band besitzt.

Auf Grund des Satzes 1 werden wir für die Häufungswerte der relativen Häufigkeiten kurz den Ausdruck " Grenzhäufigkeit " gebrauchen und insbesondere von der oberen und unteren Grenzhäufigkeit sprechen.

§ 2. Das Rechnen mit W - Bändern.

Für unsere weiteren Untersuchungen benötigen wir einige Rechenregeln über die W-Bänder, die sich ohne weiteres aus

---

2) vgl.E.Kamke, Einführung in die Wahrscheinlichkeitstheorie 1932, S.8.

den nachstehenden Hilfssätzen über den unteren und oberen Limes beschränkter Zahlenfolgen ergeben.

Es seien a_1 , a_2 , a_3 , ... und b_1 , b_2 , b_3 , ... zwei beschränkte Zahlenfolgen. Für unsere Zwecke genügt es, die erforderlichen Hilfssätze für die Voraussetzung aufzustellen, dass für schliesslich alle n

$0 \leqq a_n \leqq 1$ und $0 \leqq b_n \leqq 1$

gelten sollen. Wir bezeichnen :

$g_1 = \underline{\lim}_{n=\infty} a_n$; $G_1 = \overline{\lim}_{n=\infty} a_n$;

$g_2 = \underline{\lim}_{n=\infty} b_n$; $G_2 = \overline{\lim}_{n=\infty} b_n$.

Hilfssatz 1. Ist für schliesslich alle n $a_n \leqq b_n$, so gilt

$g_1 \leqq g_2$ und $G_1 \leqq G_2$.

Beweis. Zu $\epsilon > 0$ gibt es ein N derart, dass für <u>alle</u> $n \geqq N$

$a_n \leqq b_n$ und $a_n \geqq g_1 - \epsilon$ gleichzeitig erfüllt ist.

Für unendlich viele n gilt $b_n \leqq g_2 + \epsilon$.

Für diese unendlich viele n gilt dann :

$g_1 - \epsilon \leqq a_n \leqq b_n \leqq g_2 + \epsilon$, also

$g_1 \leqq g_2 + 2\epsilon$.

Entsprechend wird gezeigt :

$G_1 \leqq G_2 + 2\epsilon$.

Für $\epsilon \to 0$ folgt die Behauptung.

Hilfssatz 2. Ist $\underline{\lim}_{n=\infty} (a_n + b_n) = g$ und $\overline{\lim}_{n=\infty} (a_n + b_n) = G$,

so gilt $g_1 + g_2 \leqq g \leqq G \leqq G_1 + G_2$.

Beweis. Zu $\epsilon > 0$ gibt es ein N derart, dass für alle $n \geqq N$

$a_n \geqq g_1 - \epsilon$ und $b_n \geqq g_2 - \epsilon$ gleichzeitig erfüllt ist.

Für solche n ist $a_n + b_n \geqq g_1 + g_2 - 2\epsilon$.

Entsprechend wird $a_n + b_n \leqq G_1 + G_2 + 2\epsilon$.

Da die Ungleichungen für schliesslich alle n gäten, so wird auch :

$g_1 + g_2 - 2\epsilon \leqq g \leqq G \leqq G_1 + G_2 + 2\epsilon$.

Daraus folgt für $\epsilon \to 0$ die Behauptung.

Hilfssatz 3. Ist $\underline{\lim}_{n=\infty} (a_n - b_n) = g$ und $\overline{\lim}_{n=\infty} (a_n - b_n) = G$,

so gilt $g_1 - G_2 \leqq g \leqq G \leqq G_1 - g_2$.

Beweis. Zu $\epsilon > 0$ gibt es ein N derart, dass für alle $n \geqq N$

$a_n \leqq G_1 + \epsilon$ und $b_n \geqq g_2 - \epsilon$ gleichzeitig erfüllt ist.

Für solche n ist $a_n - b_n \leqq G_1 - g_2 + 2\epsilon$.

Entsprechend wird $a_n - b_n \geqq g_1 - G_2 - 2\epsilon$.

Da die Ungleichungen für schliesslich alle n gelten, so wird auch :

$g_1 - G_2 - 2\epsilon \leqq g \leqq G \leqq G_1 - g_2 + 2\epsilon$.

Daraus folgt für $\epsilon \to 0$ die Behauptung.

Hilfssatz 4. Ist $\underline{\lim}_{n=\infty} a_n \cdot b_n = g$ und $\overline{\lim}_{n=\infty} a_n \cdot b_n = G$,

so gilt $g_1 g_2 \leqq g \leqq G \leqq G_1 G_2$.

Beweis.

1. Ist $g_1 g_2 > 0$, so gibt es zu $\epsilon > 0$ mit $\epsilon < g_1$, $\epsilon < g_2$ ein N derart, dass für alle $n \geqq N$ $a_n \leqq G_1 + \epsilon$ und $b_n \leqq G_2 + \epsilon$ gleichzeitig erfüllt ist. Für solche n ist $a_n b_n \leqq (G_1 + \epsilon) (G_2 + \epsilon)$. Entsprechend wird $a_n b_n \geqq (g_1 - \epsilon) (g_2 - \epsilon)$ gezeigt. Da die Ungleichungen für schliesslich alle n gelten, so wird auch
$(g_1 - \epsilon) (g_2 - \epsilon) \leqq g \leqq G \leqq (G_1 + \epsilon) (G_2 + \epsilon)$.
Daraus folgt für $\epsilon \to 0$ die Behauptung.
2. Ist $g_1 g_2 = 0$, so ist die Behauptung trivial, da wegen $a_n = 0$ und $b_n = 0$ auch $g = 0$ ist.

Eine Verschärfung der Behauptungen durch Weglassen der Gleichheitszeichen ist nicht möglich. Es kann vorkommen, dass es zu $\epsilon > 0$ unendlich viele Zahlen n gibt, für die $a_n < g_1 + \epsilon$, und für welche gleichzeitig auch $b_n < g_2 + \epsilon$ ist, sodass in einem solchen Fall z.B.
$\underline{\lim} (a_n + b_n) = \underline{\lim} a_n + \underline{\lim} b_n$ ist.

Aus diesen Hilfssätzen folgen nun die Rechenregeln. Um diese möglichst einfach aussprechen zu können, bezeichnen wir das Intervall $g_i \leqq w \leqq G_i$ mit B_i .

Es bedeute dann

$B_1 < B_2$, dass $g_2 \leqq g_1 \leqq G_1 \leqq G_2$ ist,

$B_1 + B_2$ das Intervall $g_1 + g_2 \leqq w \leqq G_1 + G_2$,

$B_1 \cdot B_2$ das Intervall $g_1 g_2 \leqq w \leqq G_1 G_2$.

Für den Fall dass $g_1 + g_2$ oder $G_1 + G_2$ Werte annehmen, die grösser als 1 sind, werden diese Zahlen durch 1 ersetzt.

Bei diesen Bezeichnungen gelten folgende Regeln :

(a) Aus $B_1 < B_2$ und $B_2 < B_3$ folgt $B_1 < B_3$.

(b) $B_1 + B_2 = B_2 + B_1$

(c) $B_1 + (B_2 + B_3) = (B_1 + B_2) + B_3$

(d) $B_1 \cdot B_2 = B_2 \cdot B_1$

(e) $B_1 (B_2 \cdot B_3) = (B_1 \cdot B_2) B_3$

(f) $B_1 (B_2 + B_3) = B_1 B_2 + B_1 B_3$

(g) Aus $B_1 \subset B_2$ und $B_3 \subset B_4$ folgt $B_1 + B_3 \subset B_2 + B_4$

(h) Aus $B_1 \subset B_2$ und $B_3 \subset B_4$ folgt $B_1 \cdot B_3 \subset B_2 \cdot B_4$

§ 3. Unempfindlichkeitssatz.

In einer E-Folge ändert man gewisse Elemente E_{ν_1} , E_{ν_2} , E_{ν_3} , ... willkürlich ab oder lässt solche Elemente ganz weg. Die Frage, welche maximale Abweichung des W-Bandes durch diese Streichungen oder Abänderungen bewirkt wird, beantwortet der folgende Satz.

Satz 2. Ist $E : E_1 , E_2 , E_3 , \ldots$ eine E-Folge mit dem W-Band $\left[g,G \right]$ und $E'' : E_{\nu_1} , E_{\nu_2} , E_{\nu_3} , \ldots$ eine Teilfolge von E , die in E mit der oberen Grenzhäufigkeit $\tau < 1$ vorkommt, so gilt für das W-Band $\left[g^*,G^* \right]$ der Folge E^*, die aus E durch Streichung oder Abänderung der Elemente der Teilfolge E'' hervorgeht :

$$g - \tau \leqq g^* \leqq G^* \leqq \frac{G + \tau}{1 - \tau} \quad .$$

Der Satz 2 lässt sich in einer schwächeren Form aussprechen. Es ist nämlich

$$\frac{G + \tau}{1 - \tau} = G + \frac{\tau}{1 - \tau} (1 + G) .$$

Diese obere Schranke interessiert uns nur solange, als sie nicht grösser als 1 ist. Dann muss $\frac{G + \tau}{1 - \tau} \leqq 1$ oder

$G \leqq 1 - 2\tau$ sein. Daraus folgt $\frac{\tau}{1 - \tau} (1 + G) \leqq 2\tau$.

Man kann also dem Satz 2 die folgende, schwächere Form geben :

Satz 2a . Streichungen oder Abänderungen, welche mit der oberen Grenzhäufigkeit τ vorgenommen werden, verursachen eine Verbreiterung des W-Bandes um höchstens 2τ nach jeder Seite.

In diesem Satz ist als spezieller Fall ($\tau = 0$) der Unempfindlichkeitssatz enthalten :

Satz 2b . Streichungen oder Abänderungen, welche mit der Grenzhäufigkeit 0 vorgenommen werden, verbreitern das W-Band nicht.

Die Bedeutung des Unempfindlichkeitssatzes liegt für die Anwendung der Wahrscheinlichkeitsrechnung z.B. darin, dass bei einer durch Beobachtungen gewonnenen Reihe von Ergebnissen etwaige Beobachtungsfehler sicher dann als belanglos angesehen werden können, wenn sie so selten vorkommen, dass sie in einer idealisierten Folge die Grenzhäufigkeit 0 haben würden. Kommen solche Fehler mit der oberen Grenzhäufigkeit τ vor, so gibt Satz 2 oder Satz 2a die Grösse des dadurch bedingten möglichen Fehlers an.

Nun zum Beweis von Satz 2 !

Mit E' wird die zu E" komplementäre Teilfolge von E bezeichnet. Betrachtet man die n ersten Elemente von E , so kommen unter diesen n' Elemente von E' und n" Elemente von E" vor; es ist stets $n' + n'' = n$; n' und n" sind durch n eindeutig bestimmt.

Es werde n^* so gewählt, dass unter den n^* ersten Elementen von E^* genau n' Elemente von E' vorkommen. Für die Folge E^* gilt, da sicher $n^* \leq n' + n''$ ist, somit

$n' \leq n^* \leq n$.

Nach Voraussetzung ist

$$\overline{\lim_{n\to\infty}} \frac{n''}{n} = \tau \ .$$

Nun ist

$$0 \leq \left| H_{n^*}(E^*) - H_n(E) \right| \leq n'' \ .$$

Dividiert man hier durch n, so ergibt sich

$$h_n(E) - \frac{n''}{n} \leq h_{n^*}(E^*)\ \frac{n^*}{n} \leq h_n(E) + \frac{n''}{n} \ .$$

$$\left(h_n(E) - \frac{n''}{n} \right) \frac{n}{n^*} \leq h_{n^*}(E^*)$$

$$h_{n^*}(E^*) \leq \left(h_n(E) + \frac{n''}{n} \right) \frac{n}{n^*} \ .$$

Wegen $\frac{n}{n^*} \leq \frac{n}{n'}$ und $\frac{n}{n^*} \geq 1$ folgt dann :

$$h_n(E) - \frac{n''}{n} \leq h_n^*(E^*) \leq \left(h_n(E) + \frac{n''}{n}\right)\frac{n}{n'} .$$

Da nun $\frac{n}{n'} = \frac{1}{1 - \frac{n''}{n}}$

und also $\overline{\lim_{n=\infty}} \frac{n}{n'} \leq \frac{1}{1-\tau}$ ist,

so ergibt sich nach den Hilfssätzen des § 2 :

$$g - \tau \leq g^* \leq G^* \leq \frac{G + \tau}{1 - \tau} .$$

§ 4. Additionssätze.

Wenn m_1 , m_2 , m_3 , ... endlich oder abzählbar unendlich viele einander paarweise ausschliessende Merkmale sind, so versteht man unter der Merkmalmischung $m = m_1 + m_2 + m_3 + \ldots$ das Merkmal, dass auf ein Element der E-Folge entweder m_1 oder m_2 oder m_3 oder ... zutrifft.

Setzt man voraus, dass die Anzahl der Merkmale endlich, etwa gleich k ist, so gilt für jedes n, da die Merkmale sich gegenseitig ausschliessen :

$$H_n(E,m) = H_n(E,m_1) + H_n(E,m_2) + \ldots + H_n(E,m_k)$$

$$h_n(E,m) = h_n(E,m_1) + h_n(E,m_2) + \ldots + h_n(E,m_k)$$

Daraus folgt nach Hilfssatz 2 :

$g \geqq g_1 + g_2 + \dots + g_k$ und $G \leqq G_1 + G_2 + \dots + G_k$.

Es lässt sich also folgender Additionssatz aussprechen :

Satz 3. Hat eine E-Folge in Bezug auf die endlich vielen einander paarweise ausschliessenden Merkmale m_1 , m_2 , ... , m_k die W-Bänder $B_1 = B (E,m_1)$, $B_2 = B (E,m_2)$, ... , $B_k = B (E,m_k)$, so hat die gegebene E-Folge in Bezug auf $m = m_1 + m_2 + \dots + m_k$ ein W-Band $B = B (E,m)$ und es gilt :

$B \subset B_1 + B_2 + \dots + B_k$.

Eine wesentliche Voraussetzung dieses Satzes ist, dass die Anzahl der Merkmale endlich ist. Betrachtet man[3] nämlich z.B. die Folge aller natürlichen Zahlen und bedeutet m_k das Auftreten der Zahl k ($k = 1 , 2 , 3 , \dots$), so gilt für alle k $g_k = G_k = 0$ und also

$B_1 + B_2 + B_3 + \dots = [0 , 0]$.

Andererseits ist für alle n

$h_n (E,m_1 + m_2 + m_3 + \dots) = 1$.

Für unendlich viele Merkmale m_1 , m_2 , m_3 , ... gilt aber wenigstens die Ungleichung am linken Ende des Intervalls.

---

3) vgl. E.Kamke, a.a.O.,S.9.

Für eine endliche Zahl s ist sicher

$$H_n(E,m) \geqq H_n(E,m_1) + H_n(E,m_2) + \cdots + H_n(E,m_s)$$

$$h_n(E,m) \geqq h_n(E,m_1) + h_n(E,m_2) + \cdots + h_n(E,m_s).$$

Bei festem s folgt daraus nach den Hilfssätzen des § 2 :

(1) $$g \geqq g_1 + g_2 + \cdots + g_s$$

Da $g \leqq 1$ ist, so konvergiert die Reihe

$$g_1 + g_2 + g_3 + \cdots \quad .^{4)}$$

Die Teilsummen sind beschränkt und wachsen monoton. Die Beziehung (1) gilt für alle s , also ist auch

$$g \geqq g_1 + g_2 + g_3 + \cdots \quad .$$

Dieses Ergebnis lässt sich nun mit Satz 3 zusammenfassen :

Satz 4. Hat eine E-Folge E in Bezug auf die endlich oder abzählbar unendlich vielen einander ausschliessenden Merkmale

$m_1 , m_2 , m_3 , \cdots$ die W-Bänder

$\left[g_1 , G_1 \right] , \left[g_2 , G_2 \right] , \left[g_3 , G_3 \right] , \cdots ,$

so hat die E-Folge in Bezug auf die Merkmalmischung

$m = m_1 + m_2 + m_3 + \cdots$ ein W-Band $\left[g , G \right]$ und es ist stets :

$$g \geqq g_1 + g_2 + g_3 + \cdots$$

Nur dann, wenn die Anzahl der Merkmale endlich ist, gilt ausserdem

$$G \leqq G_1 + G_2 + G_3 + \cdots + G_k \quad .$$

4.) vgl. K.Knopp, Theorie und Anwendung der unendlichen Reihen, S.77 unten.

§ 5. Der V_k - Bereich .

Wir betrachten jetzt den Fall, dass auf jedes Element der E-Folge E eines und nur eines der endlich vielen Merkmale $m_1 , m_2 , \dots , m_k$ zutrifft. Da in diesem Fall stets $h_n (E,m_1) + h_n (E,m_2) + \dots + h_n (E,m_k) = 1$ ist, so müssen zwischen den Grenzhäufigkeiten in Bezug auf die Merkmale gewisse Beziehungen bestehen. Ist für $i = 1 , 2 , \dots , k$ $\quad B (E,m_i) = \left[g_i , G_i \right]$, so gibt es zu jedem w_1 in $g_1 \leq w_1 \leq G_1$ eine unendliche Teilfolge $\left\{ n_\nu \right\}$, für welche $\left\{ h_{n_\nu} (E,m_1) \right\}$ gegen w_1 konvergiert. Die Folge $\left\{ h_{n_\nu} (E,m_2) \right\}$ besitzt einen oder mehrere Häufungswerte. Man kann also wieder eine unendliche konvergente Teilfolge auswählen. Setzt man das Verfahren fort, so gelangt man zu einer unendlichen Teilfolge $\left\{ p_\nu \right\}$ derart, dass $h_{p_\nu} (E,m_i)$ für $\nu \to \infty$ gleichzeitig für alle $i = 1 , 2 , \dots , k$ konvergieren. Für diese Teilfolge seien $w_1 , w_2 , \dots , w_k$ die Grenzwerte; dann gilt $w_1 + w_2 + \dots + w_k = 1$.

Dieser Gedankengang liegt der Definition des V_k-Bereichs zugrunde :

Def.III. Ist E eine E-Folge, auf deren Elemente stets eines und nur eines der endlich vielen Merkmale $m_1 , m_2 , \dots , m_k$ zutrifft, so versteht man unter dem Verteilungsbereich für k Merkmale oder dem V_k-Bereich die Häufungsmenge,

genauer die Menge der Limespunkte der Punktfolge

$P_n = \left\{ h_n (E,m_1) , h_n (E,m_2) , \ldots , h_n (E,m_k) \right\}$

im k-dimensionalen Raum Σ_k .

V_k ist sicher eine Teilmenge des Durchschnitts D des Parallelflachs:

$g_i \leq w_i \leq G_i \qquad (i = 1 , 2 , 3 , .., k)$

und der Ueberebene

$w_1 + w_2 + \ldots + w_k = 1$.

Dabei sind $w_1 , w_2 , \ldots , w_k$ die Koordinaten der Punkte von V_k .

Im Fall $k = 2$ ist $V_k = D$. Da stets $w_2 = 1 - w_1$ ist, so durchläuft w_2 die Werte von $1 - g_1$ bis $1 - G_1$, wenn w_1 die Werte von g_1 bis G_1 durchläuft. V_k ist also die Strecke, die die Punkte $\left[g_1 , 1 - g_1 \right]$ und $\left[G_1 , 1 - G_1 \right]$ verbindet.

Ist $k > 2$, so kann V_k eine echte Teilmenge des Durchschnitts D sein. Dies ist in dem folgenden Beispiel der Fall :

Die aus den Merkmalen 0 , 1 und 2 aufgebaute E-Folge wird in Abschnitte eingeteilt. Der (m+1)te Abschnitt hat die Länge $3 \cdot 2^m$; auf 2^m mal 0 folgt 2^m mal 1 und dann 2^m mal 2 . Die Folge beginnt also

0,1,2 ; 00,11,22 ; 0000,1111,2222 ; ...

Setzt man $n = 3(2^m - 1) + \lambda$ mit $0 < \lambda \leq 3 \cdot 2^m$, so bilden die Elemente E_n den (m+1)ten Abschnitt.

Wir betrachten nun den Teilabschnitt $0 \leq \lambda \leq 2^m$ und setzen $\alpha' = \frac{\lambda}{2^m}$. Es ist nun, wenn m_i für $i = 1, 2, 3$ das Merkmal 0, 1, 2 bedeutet

$$h_n(E, m_1) = \frac{2^m - 1 + \lambda}{3(2^m - 1) + \lambda} = \frac{(1 + \alpha')2^m - 1}{(3 + \alpha')2^m - 3}$$

$$h_n(E, m_2) = \frac{2^m - 1}{3(2^m - 1) + \lambda} = \frac{2^m - 1}{(3 + \alpha')2^m - 3}$$

$$h_n(E, m_3) = \frac{2^m - 1}{3(2^m - 1) + \lambda} = \frac{2^m - 1}{(3 + \alpha')2^m - 3}$$

Für festes $0 \leq \alpha \leq 1$ wählt man zu jedem m ein α', das möglichst nahe an α liegt; dies ist stets möglich, da mit wachsendem m die α' das Intervall von 0 bis 1 beliebig dicht erfüllen. Für $m \to \infty$ geht man zu den Grenzwerten über :

$$w_1 = \frac{1 + \alpha}{3 + \alpha} \quad ; \quad w_2 = \frac{1}{3 + \alpha} \quad ; \quad w_3 = \frac{1}{3 + \alpha}$$

Führt man jetzt $\alpha = \frac{3t}{4 - t}$ ein, so durchläuft t alle Werte von 0 bis 1, wenn α die Werte von 0 bis 1 durchläuft,

da $t = \frac{4\alpha}{3+\alpha}$ und $\frac{dt}{d\alpha} = \frac{12}{(3+\alpha)^2} > 0$ ist.

Es ergibt sich dann :

$$w_1 = \frac{1}{6}t + \frac{1}{3} \quad ; \quad w_2 = -\frac{1}{12}t + \frac{1}{3} \quad ; \quad w_3 = -\frac{1}{12}t + \frac{1}{3} \quad .$$

Entsprechend findet man für $2^m \leqq \lambda \leqq 2 \cdot 2^m$

$$w_1 = -\frac{1}{10}t + \frac{1}{2} \quad ; \quad w_2 = \frac{3}{20}t + \frac{1}{4} \quad ; \quad w_3 = -\frac{1}{20}t + \frac{1}{4}$$

und für $2 \cdot 2^m \leqq \lambda \leqq 3 \cdot 2^m$

$$w_1 = -\frac{1}{15}t + \frac{2}{5} \quad ; \quad w_2 = -\frac{1}{15}t + \frac{2}{5} \quad ; \quad w_3 = \frac{2}{15}t + \frac{1}{5} \quad .$$

V_k besteht also aus den drei Strecken, die die drei Punkte

$$\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \quad ; \quad \frac{1}{2}, \frac{1}{4}, \frac{1}{4} \quad ; \quad \frac{2}{5}, \frac{2}{5}, \frac{1}{5}$$

verbinden; nicht zu V_k gehören die Punkte im Innern des Dreiecks. V_k braucht also nicht konvex zu sein; jedoch gilt, wie wir für beliebiges k jetzt noch zeigen wollen, dass V_k zusammenhängend ist.[5)]

Man betrachtet zwei Punkte P_n und P_{n+1} . Diese haben die

Entfernung $\rho = \sqrt{\sum_{i=1}^{k} \left(h_{n+1}(E,m_i) - h_n(E,m_i) \right)^2}$

---

5) vgl. F.Hausdorff, Grundzüge der Mengenlehre 1935, S. 130, 158, 159 .

Nun ist $H_{n+1}(E,m_i) = H_n(E,m_i) + \vartheta$, wo ϑ entweder 0 oder 1 ist. Dividiert man durch n, so ergibt sich

$$h_{n+1}(E,m_i) - h_n(E,m_i) = \frac{\vartheta - h_{n+1}(E,m_i)}{n}$$

Da die rechte Seite nur Werte zwischen $-\frac{1}{n}$ und $+\frac{1}{n}$ annehmen kann, so gilt :

$$\left| h_{n+1}(E,m_i) - h_n(E,m_i) \right| \leq \frac{1}{n}$$

Ist $\epsilon > 0$ gegeben und wählt man $N > \frac{\sqrt{k}}{\epsilon}$, so ist für alle $n \geq N$:

$$\vartheta < \sqrt{\sum_{i=1}^{k} \frac{\epsilon^2}{k}} = \epsilon .$$

Zu jedem Punkt P von V_k bildet man die (offene) Umgebung U_ϵ (P) ; dabei ist U_ϵ (P) die Menge aller Punkte, die von P einen Abstand kleiner als ϵ haben. Durch das System dieser Umgebungen wird die abgeschlossene Menge V_k überdeckt. Nach dem Borel'schen Ueberdeckungssatz genügen schon endlich viele solcher Umgebungen U_ϵ' , U_ϵ'' , ... , $U_\epsilon^{(r)}$ zur Ueberdeckung von V_k. Es gibt nun, da die $\left\{ P_n \right\}$ beschränkt sind, eine Zahl $N_1 \geq N$ derart, dass P_n für $n \geq N_1$ in einer solchen Umgebung liegt; andererseits liegen in jeder $U_\epsilon^{(\rho)}$ unendlich viele P_n , da ja in jeder $U_\epsilon^{(\rho)}$ mindestens ein Punkt von V_k liegt.

Es seien nun Q' und Q" zwei beliebige Punkte von V_k . Q' sei etwa in U_ε' , Q" etwa in U_ε'' enthalten. P_{n_o} mit $n_o > N_1$ liege in U_ε' . In der Punktfolge P_{n_o} , P_{n_o+1} , P_{n_o+2} , ... kommt sicher ein Punkt, etwa P_{n_o+p} , vor, der in U_ε'' enthalten ist. Jeder Punkt P_{n_o+i} mit $0 \leq i \leq p$ gehört einer Umgebung $U_\varepsilon^{(\rho)}$ an, in der sicher mindestens ein Punkt Q_i von V_k enthalten ist; $Q_0 = Q'$, $Q_p = Q''$. Diesen Punkt Q_i ordnen wir dem Punkt P_{n_o+i} zu; dass verschiedenen Punkten P_{n_o+i} der gleiche Punkt Q zugeordnet sein kann, stört nicht. Zwei aufeinanderfolgende Punkte Q_i und Q_{i+1} haben die Entfernung

$$\overline{Q_iQ_{i+1}} \leq \overline{Q_iP_{n_o+i}} + \overline{P_{n_o+i}P_{n_o+i+1}} + \overline{P_{n_o+i+1}Q_{i+1}} \text{ , also}$$

$$\overline{Q_iQ_{i+1}} < 2\varepsilon + \varepsilon + 2\varepsilon = 5\varepsilon .$$

Die Punkte Q_0 , Q_1 , ... , Q_p bilden eine ρ-Kette mit $\rho < 5\varepsilon$. Da ε beliebig klein gewählt werden kann, so ist die Distanz $\underline{\lim}\ \rho = 0$: Die abgeschlossene Menge V_k ist zusammenhängend.

§ 6. X - Folgen.

Wir betrachten wieder alternierende E-Folgen mit den Merkmalen m und $\overline{m}$. Ueber die E-Folgen wurde bisher lediglich vorausgesetzt, dass sie unendliche Folgen sein sollen. Wir unterwerfen nun die E-Folgen einer ersten Einschränkung und erhalten so die Menge aller X-Folgen, die eine Teilmenge der Menge aller E-Folgen ist.

Es sei r und k fest gewählt und $0 \leq r \leq k$, $k \geq 1$. Die E-Folge mit dem W-Band $\left[g,G\right]$ schreiben wir nun (k+1)-mal untereinander an und zwar so, dass wir das erste Mal mit E_1 , das zweite Mal mit E_2 , ... , das (k+1)-te Mal mit E_{k+1} beginnen :

E_1 , E_2 , E_3 , ... , E_n , ...
E_2 , E_3 , E_4 , ... , E_{n+1} , ...
.....
E_k , E_{k+1} , E_{k+2} , ... , E_{n+k-1} , ...
E_{k+1} , E_{k+2} , E_{k+3} , ... , E_{n+k} , ...

Dieses Schema bezeichnen wir als $F^{(k+1)}$-Feld. Die einzelnen Zeilen des Feldes werden der Reihe nach mit E' , E'' , ... , $E^{(k)}$, $E^{(k+1)}$ bezeichnet.

Jetzt betrachten wir die n-te Spalte des Feldes $F^{(k+1)}$. Wenn unter den k Elementen E_n , E_{n+1} , E_{n+2} , ... , E_{n+k-1} genau r Elemente mit dem Merkmal m vorkommen, so sagen wir

E_{n+k} habe eine Vorgängerkombination aus r Elementen m und k-r Elementen $\bar{m}$.

Mit $E^{(k,r)}$ wird diejenige Teilfolge von $E^{(k+1)}$ bezeichnet, deren Elemente eine Vorgängerkombination aus r Elementen m und (k-r) Elementen $\bar{m}$ haben. $E^{(k,r)}$ braucht nun, wenn keine weiteren Voraussetzungen über die E-Folge gemacht werden, keine unendliche Folge zu sein. Verlangt man dies aber von der E-Folge, so kommt man zu den X-Folgen. Deren genaue Definition kann so ausgesprochen werden :

Def.IV. Unter einer X-Folge wird eine E-Folge verstanden, die folgender Bedingung genügt :
Für alle $r = 0, 1, \dots, k$ und für alle $k = 1, 2, 3, \dots$ ist die Folge $E^{(k,r)}$ eine unendliche Folge.

Ist $B^{(k,r)}$ das W-Band von $E^{(k,r)}$, so bezeichne $B^{*} = \left[g^{*}, G^{*} \right]$ das kleinste Intervall, in dem B(E,m) und alle $B^{(k,r)}$ enthalten sind. Es ist offenbar

(*) $\quad g^{*} \leq g \leq G \leq G^{*}$

Man betrachtet die Folge $E^{(k+1)}$. Wenn ein Element dieser Folge eine Vorgängerkombination aus r Merkmalen m und (k-r) Merkmalen $\bar{m}$ besitzt, so ordnen wir diesem Element das Merkmal m_r zu. Da die so definierten Merkmale $m_0 , m_1 , \dots , m_k$ sich gegenseitig ausschliessen, so besitzt $E^{(k+1)}$ in Bezug auf $m_0 , m_1 , \dots , m_k$ einen

V_{k+1}-Bereich. Es gibt also eine Teilfolge n_1 , n_2 , n_3 , ... der Folge der natürlichen Zahlen derart, dass

$$\lim_{\nu=\infty} h_{n_\nu}(E^{(k+1)},m_r) = w(k,r)$$

existiert; dabei ist

$$\sum_{r=0}^{k} w(k,r) = 1 .$$

Der V_{k+1}-Bereich in Σ_{k+1} besteht aus Punkten

$$\{ w(k,0) , w(k,1) , \dots , w(k,k) \} .$$

Nun betrachten wir den V_k-Bereich der Folge $E^{(k)}$ und den V_{k+1}-Bereich der Folge $E^{(k+1)}$. Zwei Punkte $Q^{(k)} \prec V_k$ und $Q^{(k+1)} \prec V_{k+1}$ heissen zugeordnet, wenn es eine Teilfolge n_1 , n_2 , n_3 , ... der Folge der natürlichen Zahlen (vgl. § 5) gibt, sodass

$$\lim P \{ h_{n_\nu}(E^{(k)},m_r^{(k)}) \} = Q^{(k)} \quad \text{und}$$

$$\lim P \{ h_{n_\nu}(E^{(k+1)},m_r^{(k+1)}) \} = Q^{(k+1)}$$

gleichzeitig existieren. (Der obere Index bei m deutet die Länge der Vorgängerkombination an.)

Zu jedem $Q^{(k)}$ von V_k gibt es (mindestens) einen zugeordneten Punkt $Q^{(k+1)}$ von V_{k+1} . Denn wenn $Q^{(k)}$ Punkt von V_k ist, so gibt es stets eine Folge n_1' , n_2' , n_3' , ... derart, dass $w(k-1,s) = \lim_{\nu=\infty} h_{n_\nu'}(E^{(k)},m_s^{(k)})$ gleichzeitig für alle $s = 0, 1, \dots , k-1$ gilt. Die Folgen $h_{n_\nu'}(E^{(k+1)},m_r^{(k+1)})$ sind für $r = 0, 1, \dots , k$

beschränkte Zahlenfolgen; es gibt also mindestens eine Teilfolge $n_1, n_2, n_3, \ldots$ der Folge $n'_1, n'_2, n'_3, \ldots$ derart, dass $\lim_{\nu=\infty} h_{n_\nu}(E^{(k+1)}, m_r^{(k+1)})$ existiert für $r = 0, 1, \ldots, k$.

Bezeichnet man diese Grenzwerte mit $w(k,r)$, so ist

$$Q^{(k+1)} = \{ w(k,0), w(k,1), \ldots, w(k,k) \}$$

also ein Punkt von V_{k+1} und offenbar $Q^{(k)}$ zugeordnet, da ja auch $\lim_{\nu=\infty} h_{n_\nu}(E^{(k)}, m_s^{(k)}) = w(k-1,s)$ ist.

Nun soll eine Beziehung zwischen den Koordinaten $w(k-1,r)$ und den Koordinaten $w(k,r)$ aufgestellt werden. Dazu führen wir zwei sich ausschliessende Merkmale $M'_{k,r}$ und $M''_{k,r}$ ein :

$M'_{k,r}$ bedeute die Vorgängerkombination aus r Elementen m und (k-r) Elementen $\bar{m}$, wobei das letzte Element m sei.

$M''_{k,r}$ bedeute die Vorgängerkombination aus r Elementen m und (k-r) Elementen $\bar{m}$, wobei das letzte Element $\bar{m}$ sei.

Deuten wir die Merkmalmischung symbolisch durch das Plus-Zeichen an, so gilt :

$$m_r^{(k)} = M'_{k,r} + M''_{k,r} \quad ;$$

$$m_r^{(k-1)} = M'_{k,r+1} + M''_{k,r} \quad \text{und somit :}$$

(1) $$H_n(E^{(k+1)}, m_r^{(k)}) = H_n(E^{(k+1)}, M'_{k,r}) + H_n(E^{(k+1)}, M''_{k,r})$$

(2) $$H_n(E^{(k)}, m_r^{(k-1)}) = H_n(E^{(k+1)}, M'_{k,r+1}) + H_n(E^{(k+1)}, M''_{k,r})$$

Da nach Def.IV $H_n(E^{(k)}, m_r^{(k-1)}) > 0$ für schliesslich

alle n ist, so kann man setzen :

(3) $H_n(E^{(k+1)}, M'_{k,r+1}) = a_n^{(k,r)} \cdot H_n(E^{(k)}, m_r^{(k-1)})$

Nach (2) wird dann :

(4) $H_n(E^{(k+1)}, M''_{k,r}) = (1-a_n^{(k,r)}) H_n(E^{(k)}, m_r^{(k)})$

Ersetzt man in (3) r durch r-1 , so folgt aus (1), (3) und (4) :

(5) $h_n(E^{(k+1)}, m_r^{(k)}) = a_n^{(k,r-1)} h_n(E^{(k)}, m_{r-1}^{(k-1)})$

$+ (1-a_n^{(k,r)}) h_n(E^{(k)}, m_r^{(k-1)})$

$a_n^{(k,r)}$ ist die Anzahl derjenigen Elemente von $E^{(k+1)}$, welche eine Vorgängerkombination aus r+1 Elementen m und k-r-1 Elementen $\bar{m}$ haben, wobei das letzte Element m ist, bezogen auf die Anzahl der Elemente von $E^{(k)}$, die eine Vorgängerkombination aus r Elementen m und k-r-1 Elementen $\bar{m}$ haben. Das ist aber die Häufigkeit von m in $E^{(k-1,r)}$ bezogen auf die Anzahl $n' = H_n(E^{(k)}, m_r)$, also $a_n^{(k,r)} = h_{n'}(E^{(k-1,r)}, m)$ mit $n' = H_n(E^{(k)}, m_r)$. Nach der Definition des Intervalls $[g^*; G^*]$ gilt dann für $n > N(\epsilon)$

$g^* - \epsilon \leq a_n^{(k,r)} \leq G^* + \epsilon$.

Setzt man nun voraus, dass $n'_1, n'_2, n'_3, \ldots$ eine Folge von wachsenden natürlichen Zahlen ist, für die

$\lim_{\nu=\infty} h_{n'_\nu}(E^{(k)}, m_r^{(k-1)}) = w(k-1,r)$

für alle $r = 0, 1, \ldots, k$; $k \geq 1$ existieren, so

bilden die Zahlenfolgen $a_{n_\nu}^{(k,r)}$ beschränkte und wegen der Definitionseigenschaft der X-Folgen unendliche Folgen; durch weitere Teilfolgenauswahl $n_1, n_2, n_3, \ldots$ lässt sich erreichen, dass auch

$$\lim_{\nu=\infty} a_{n_\nu}^{(k,r)} = U(k,r) \quad \text{für} \quad r = 0, 1, \ldots, k$$

existieren. Für diese Teilfolge existieren nach (5) auch die Grenzwerte

$$\lim_{\nu=\infty} h_{n_\nu}(E^{(k+1)}, m_r^{(k)}) = w(k,r) \qquad \text{und es gilt}$$

(6) $$w(k,r) = (1-U(k,r))\, w(k-1,r) + U(k,r-1)\, w(k-1,r-1)$$

wobei

(7) $$g^* \leq U(k,r) \leq G^* \quad \text{ist.}$$

Die Formel (6) wird noch durch die plausiblen Festsetzungen $w(k,-1) = w(k,k+1) = 0$ ergänzt.

Um uns kurz ausdrücken zu können, wollen wir ein System von Zahlen $w(k,r)$ $(k = 1, 2, 3, \ldots ; r = 0, 1, \ldots, k)$ ein Verteilungssystem nennen, wenn es Zahlen $U(k,r)$ gibt, sodass die folgenden Beziehungen gelten :

(I) $$w(0,0) = 1$$

(II) $$w(k,r) = (1-U(k,r))\, w(k-1,r) + U(k,r-1)\, w(k-1,r-1)$$

$$w(k,-1) = w(k,k+1) = 0$$

(III) $$a \leq U(k,r) \leq b \; ; \; 0 \leq a \leq b \leq 1 \; .$$

$[a,b]$ ist das Variationsintervall des Verteilungssystems.

Die vorausgehenden Betrachtungen lehren, dass man durch fortgesetzte Teilfolgenauswahl und schliesslich Diagonalfolgenbildung Folgen wachsender natürlicher Zahlen $n_1 , n_2 , n_3 , \ldots$ bilden kann, für welche alle $\left\{ a_{n_\nu}^{(k,r)} \right\}$ und $\left\{ h_{n_\nu}(E^{(k+1)}, m_r^{(k)}) \right\}$ konvergieren, deren Grenzwerte also ein Verteilungssystem darstellen. Die Menge der Verteilungssysteme, die man aus (I), (II), (III) für $a = g^*$, $b = G^*$ gewinnt, umfasst die Menge der aus der betrachteten X-Folge in der obigen Weise ableitbaren Verteilungssysteme.

In den beiden folgenden §§ 7 und 8 werden wir die 1. und 2. Bernoullische Aussage für Verteilungssysteme formulieren. Bemerkenswert ist dabei, dass unsere Untersuchungen sich ausschliesslich auf die Beziehungen (I), (II), (III) stützen; von irgendwelchen weiteren Eigenschaften der X-Folgen wird kein Gebrauch gemacht. In § 9 werden wir dann die Ergebnisse auch für X-Folgen formulieren.

Besonders einfach gestalten sich unsere Untersuchungen für einen Spezialfall des allgemeinen Verteilungssystems, den wir als " Poisson'schen Spezialfall " bezeichnen. Dazu setzen wir $U(t,r) = U(t)$ als von r unabhängig voraus. Setzt man sogar noch $U(t) = U_0$ als von t unabhängig voraus, so kommt man zum " Bernoullischen Fall " .

Für die Untersuchung des Poisson'schen Spezialfalls benötigen wir weiter unten einige Formeln, die hier abge-

leitet werden sollen.

Die Funktion

$$(8) \qquad E_t(x) = \sum_{r=0}^{t} w(t,r)\, x^r$$

bezeichnen wir als erzeugende Funktion des Poisson'schen Spezialfalls. Es gilt nämlich

$$(9) \qquad E_t(x) = \prod_{s=1}^{t} \left(1-U(s)+U(s)\, x \right) .$$

Der Beweis wird durch vollständige Induktion geführt :

Ist $t = 1$, so folgt aus (8), (II), (I)

$$E_1(x) = w(1,0) + w(1,1)\, x$$
$$= (1-U(1))\, w(0,0) + U(1)\, w(0,0)\, x$$
$$= 1 - U(1) + U(1)\, x \quad ,$$

sodass also (9) für $t = 1$ gilt; wir setzen voraus, dass (9) für $t \leq \tau$ richtig sei, und wollen zeigen, dass (9) dann auch für $t = \tau + 1$ gilt. Es ist

$$\prod_{s=1}^{\tau+1} \left(1-U(s)+U(s)\, x \right)$$

$$= \left(1-U(\tau+1)+U(\tau+1)\, x \right) \prod_{s=1}^{\tau} \left(1-U(s)+U(s)\, x \right)$$

und, da (9) für $t = \tau$ als richtig vorausgesetzt wird, weiter

$$(10) \qquad = \left(1-U(\tau+1)+U(\tau+1)\, x \right) \sum_{r=0}^{\tau} w(\tau,r)\, x^r \quad .$$

Andererseits ist wegen (II)

$$\sum_{r=0}^{\tau+1} w(\tau+1,r)\, x^r$$

$$= \sum_{r=0}^{\tau+1} x^r \Big\{ (1-U(\tau+1))\, w(\tau,r) + U(\tau+1)\, w(\tau,r-1) \Big\}$$

$$= (1-U(\tau+1)) \sum_{r=0}^{\tau} x^r w(\tau,r,) + U(\tau+1) \sum_{r=1}^{\tau+1} x^r w(\tau,r-1)$$

$$= (1-U(\tau+1)) \sum_{r=0}^{\tau} x^r w(\tau,r) + U(\tau+1)\, x \sum_{r=0}^{\tau} x^r w(\tau,r) ,$$

also weiter

(11) $$= (1-U(\tau+1)+U(\tau+1)\, x) \sum_{r=0}^{\tau} x^r w(\tau,r) \quad .$$

Aus (10) und (11) folgt die Richtigkeit von (9) für $t = \tau+1$. Damit ist der Induktionsbeweis fertig.

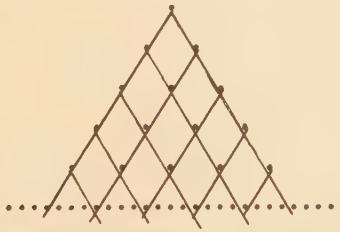
Die Verteilungssysteme lassen sich mit Hilfe des Galton'schen Brettes veranschaulichen. Wir denken uns dazu die Nägel in der üblichen Weise eingeschlagen (Fig.1) .

(Fig.1)

Sind alle Nägel senkrecht und lässt man die Masse 1 von oben einfallen, so erhält man das Verteilungssystem $g^* = G^* = 1/2$ (Bernoullischer Fall mit $w = 1/2$) . Zum allgemeinen Bernoullischen Fall gelangt man, wenn man die Nägel schief einschlägt, jedoch so, dass alle Nägel die

gleiche Richtung haben. Verlangt man nur, dass die Nägel einer und derselben Zeile die gleiche Richtung haben, dies aber für Nägel verschiedener Zeilen nicht mehr gilt, so haben wir eine Veranschaulichung des Poisson'schen Spezialfalls vor uns. Den allgemeinen Fall mit dem Variationsintervall $[a,b]$ erhält man, wenn über die Richtung der Nägel weiter nichts voraussetzt als dass die Bedingungen (III) erfüllt sind (wackeliges Galton'sches Brett) .

Die Werte $w(t,r)$ eines Verteilungssystems kann man folgendermassen deuten. In den Punkten der t-ten Zeile des Galton'schen Brettes denken wir uns die Teilmassen konzentriert, die im Galton'schen Brett über t Nagelreihen geflossen sind, wenn von oben her die Masse 1 fiel. Wir erhalten dann folgendes Schema :



Bezeichnen wir die Gerade, die durch die linken Endpunkte geht, mit $r = 0$ und die Parallelen dazu der Reihe nach mit $r = 1, 2, 3, \ldots$, so haben die Teilmassen der Zeile t die Werte $w(t,r)$; die Verbindungslinien zweier benachbarter Punkte denken wir uns mit einem Gewicht entsprechend der Schiefe des weiter oben befindlichen Nagels versehen. Diese Veranschaulichung im Galton'schen Brett wird uns gute Dienste erweisen.

§ 7. Die 1. Bernoullische Aussage für Verteilungssysteme.

1. Wir betrachten zunächst den Poisson'schen Spezialfäl.

Aus der Formel (8) des § 6 folgt durch Differentiation

(1) $$E'(x) = \sum_{r=0}^{t} r\, w(t,r)\, x^{r-1} \qquad 6)$$

$$E''(x) = \sum_{r=0}^{t} r\,(r-1)\, w(t,r)\, x^{r-2}$$

$$= \sum_{r=0}^{t} r^2\, w(t,r)\, x^{r-2} - \sum_{r=0}^{t} r\, w(t,r)\, x^{r-2}$$

(2) $$E''(x) = \sum_{r=0}^{t} r^2\, w(t,r)\, x^{r-2} - \frac{1}{x}\, E'(x) \quad (x > 0)$$

Wir bezeichnen als

Mittelwert $M(t)$ den Wert $E'(1) = \sum_{r=0}^{t} r\, w(t,r)$, als

Streuung $S(t)$ den Wert $E''(1) + E'(1) - \left\{E'(1)\right\}^2$

Es ist wegen (1) und (2) für $x = 1$ und wegen der Definition von $M(t)$:

$$S(t) = \sum_{r=0}^{t} r^2\, w(t,r) - M(t) + M(t) - \left\{M(t)\right\}^2$$

$$= \sum_{r=0}^{t} r^2\, w(t,r) - 2\left\{M(t)\right\}^2 + \left\{M(t)\right\}^2$$

$$= \sum_{r=0}^{t} r^2\, w(t,r) - 2\,M(t) \sum_{r=0}^{t} r\, w(t,r)$$

$$+ \left\{M(t)\right\}^2 \sum_{r=0}^{t} w(t,r)$$

---

6) Wir lassen hier bei E den Index t weg.

$$S(t) = \sum_{r=0}^{t} \left\{ M(t) - r \right\}^2 w(t,r) \quad .$$

Andererseits folgt aus (9) in § 6

$$E'(x) = E(x) \sum_{s=1}^{t} \frac{U(s)}{1-U(s)+U(s)\,x}$$

$$E''(x) = E'(x) \sum_{s=1}^{t} \frac{U(s)}{1-U(s)+U(s)\,x} - E(x) \sum_{s=1}^{t} \left\{ \frac{U(s)}{1-U(s)+U(s)\,x} \right\}^2$$

Für $x = 1$ wird

$$E'(1) = \sum_{s=1}^{t} U(s)$$

$$E''(1) = \left\{ E'(1) \right\}^2 - \sum_{s=1}^{t} \left\{ U(s) \right\}^2$$

und damit

(3) $$M(t) = \sum_{s=1}^{t} U(s)$$

$$S(t) = \left\{ E'(1) \right\}^2 - \sum_{s=1}^{t} \left\{ U(s) \right\}^2 + E'(1) - \left\{ E'(1) \right\}^2$$

$$S(t) = M(t) - \sum_{s=1}^{t} \left\{ U(s) \right\}^2$$

(4) $$S(t) = \sum_{s=1}^{t} U(s)\,(1-U(s))$$

Da nach § 6 (III) $a \leq U(s) \leq b$ ist, so gilt

(5) $$t\,a \leq M(t) \leq t\,b$$

(6) $$t\,a\,(1-b) \leq S(t) \leq t\,b\,(1-a) \quad .$$

Ist nun $q > 0$ mit $a > q$, $1-b > q$ fest gegeben, so wird

$$S(t) = \sum_{r=0}^{t} \{M(t) - r\}^2 w(t,r)$$

$$\geq \sum_{|r-M(t)| > t\,q} \{M(t) - r\}^2 w(t,r)$$

$$\geq t^2 q^2 \sum_{|r-M(t)| > t\,q} w(t,r)$$

Daraus folgt nun

$$\sum_{|r-M(t)| > t\,q} w(t,r) \leq \frac{1}{t^2 q^2} S(t)$$

und daraus wegen (6)

$$\sum_{|r-M(t)| > t\,q} w(t,r) \leq \frac{1}{t^2 q^2} t\,b\,(1-a) \quad \text{oder}$$

$$\sum_{|r-M(t)| > t\,q} w(t,r) \leq \frac{1}{t} \cdot \frac{b\,(1-a)}{q^2} .$$

Diese Ungleichung gilt für die Summe über alle Werte $w(t,r)$, für die $r > M(t) + t\,q$ und $r < M(t) - t\,q$ ist. Sie gilt erst recht dann, wenn $r > t\,b + t\,q$ und $r < t\,a - t\,q$, also wenn $r > t\,(b+q)$ und $r < t\,(a-q)$ ist.

2. Für $0 \leq a \leq b \leq 1$ lässt sich die 1. Bernoullische Aussage für Verteilungssysteme so aussprechen :

Bei beliebig gewähltem, dann aber festgehaltenem q mit $0 < q$ strebt der Ausdruck

$$\sum_{k(a-q) \leq r \leq k(b+q)} w(k,r)$$

mit wachsendem k gegen 1; in der Summe ist natürlich nur über solche r zu summieren, für welche $0 \leq r \leq k$ ist.

Für den Poisson'schen Fall lässt sich die Behauptung jetzt ohne weiteres beweisen; es ist nämlich

$$1 \geq \sum_{k(a-q) \leq r \leq k(b+q)} w(k,r)$$

$$= 1 - \sum_{\substack{r < k(a-q) \\ r > k(b+q)}} w(k,r)$$

$$\geq 1 - \frac{1}{k} \cdot \frac{b(1-a)}{q^2}$$

Für festes $q > 0$ und $k \to \infty$ folgt daraus die Behauptung.

Für den Bernoullischen Fall gilt $a = b = w$. Es wird dann

$$1 \geq \sum_{k(w-q) \leq r \leq k(w+q)} w(k,r) \geq 1 - \frac{1}{k} \cdot \frac{w(1-w)}{q^2} \quad .$$

Ist $a = 0$, $b < 1$, so wählt man q in $0 < q < 1-b$. Es gilt dann ebenfalls

$$\sum_{|r-M(t)| > tq} w(t,r) \leq \frac{1}{t^2 q^2} S(t) \quad .$$

Nun ist wegen (6) $S(t) \leqq t\,b$ und die Ungleichung

$$\sum_{|r-M(t)|>tq} w(t,r) \leqq \frac{1}{t} \cdot \frac{b}{q^2}$$

gilt erst recht für alle $r > t\,(b+q)$ und der Ausdruck

$$\sum_{0 \leqq r \leqq k(b+q)} w(k,r)$$ geht bei festem q mit wachsendem k gegen 1. Da die entsprechende Betrachtung auch für $a > 0$, $b = 1$ gilt und für $a = 0$, $b = 1$ der Satz trivial ist, so gilt die Behauptung auch für $0 \leqq a \leqq b \leqq 1$.

3. Nun zum allgemeinen Fall ! Dazu gehen wir wieder von den Formeln § 6 (I), (II), (III) aus :

(I) $w(0,0) = 1$

(II) $w(t,r) = (1-U(t,r))\,w(t-1,r) + U(t,r-1)\,w(t-1,r-1)$

$w(t,-1) = w(t,t+1) = 0$

(III) $a \leqq U(t,r) \leqq b \quad ; \quad 0 \leqq a \leqq b \leqq 1$.

Wir betrachten zunächst den Fall $0 < a \leqq b < 1$, k werde im folgenden zunächst festgehalten; ferner wird eine Zahl q mit $0 < a-q < b+q < 1$ gewählt und festgehalten. Nun wird für $t \leqq k$ gesetzt :

(9) $$\Omega(t,r) = \sum_{\rho=0}^{r} w(t,\rho)$$

Aus (II) folgt :

$$\Omega\ (\ t,r\ ) = \sum_{\rho=0}^{r} \Big\{ (\ 1-U(\ t,\rho\ )\ )\ w(\ t-1,\rho\ ) + U(\ t,\rho-1\ )\ w(\ t-1,\rho-1\ ) \Big\}$$

$$= \sum_{\rho=0}^{r} (\ 1-U(\ t,\rho\ )\ )\ w(\ t-1,\rho\ ) + \sum_{\rho=0}^{r-1} U(\ t,\rho\ )\ w(\ t-1,\rho\ )$$

$$= \sum_{\rho=0}^{r} w(\ t-1,\rho\ ) - U(\ t,r\ )\ w(\ t-1,r\ )$$

Mit Hilfe von (9) ergibt sich als Rekursionsformel für $\Omega\ (\ t,r\ )$:

o) $$\Omega\ (\ t,r\ ) = \Omega\ (\ t-1,r\ ) - U(\ t,r\ ) \Big\{ \Omega\ (\ t-1,r\ ) - \Omega\ (\ t-1,r-1\ ) \Big\}$$

Setzt man

$$w_a(\ t,r\ ) = \binom{t}{r} a^r\ (1-a)^{t-r}\ ;$$

$$w_b(\ t,r\ ) = \binom{t}{r} b^r\ (1-b)^{t-r}\ ;$$

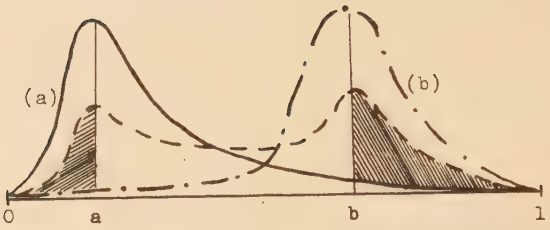
$$\Omega_a(\ t,r\ ) = \sum_{\rho=0}^{r} w_a(\ t,\rho\ )\ ;$$

$$\Omega_b(\ t,r\ ) = \sum_{\rho=0}^{r} w_b(\ t,\rho\ )\ ,$$

so gilt für $r = 0, 1, \ldots, t$

l) $$\Omega_b(\ t,r\ ) \leq \Omega\ (\ t,r\ ) \leq \Omega_a(\ t,r\ )\ .$$

(11) wird durch Schluss von $t-1$ auf t bewiesen.



Für $t = 1$ ist (11) richtig :

$$w(1,0) = \{1 - U(1,0)\} \cdot 1 + 0$$

$$w(1,1) = 0 + U(1,0) \cdot 1$$

(12)
$$\Omega(1,0) = 1 - U(1,0)$$

$$\Omega(1,1) = 1 \quad ;$$

wegen (12) und (III) ist

$$1-b \leqq \Omega(1,0) \leqq 1-a .$$

Andererseits ist

$$\Omega_a(1,0) = 1-a \; ; \; \Omega_b(1,0) = 1-b \; .$$

$$\Omega_a(1,1) = \Omega_b(1,1) = 1 \; .$$

Die Behauptung gilt also für $t = 1$; sie sei für $r \leqq t-1$ bereits bewiesen ; unter Verwendung dieser Annahme und der Beziehungen (10),(11) ergibt sich

$$\Omega(t,r) \leqq \Omega(t-1,r) - a\,(\Omega(t-1,r) - \Omega(t-1,r-1))$$
$$\leqq (1-a)\,\Omega_a(t-1,r) + a\,\Omega_a(t-1,r-1)$$
$$= \Omega_a(t,r)$$

$$\Omega(t,r) \geqq \Omega(t-1,r) - b\,(\Omega(t-1,r) - \Omega(t-1,r-1))$$
$$\geqq (1-b)\,\Omega_b(t-1,r) + b\,\Omega_b(t-1,r-1)$$
$$= \Omega_b(t,r)$$

(11) gilt also auch für $\tau = t$ und damit allgemein.

Setzt man nun

$$r_0 = \left[k\,(a-q)\right] - 1 \quad ; \quad r_1 = \left[k\,(b+q)\right] \ ,$$

so wird

(13)
$$\sum_{k(a-q) \leq r \leq k(b+q)} w(k,r) = \sum_{r=r_0+1}^{r_1} w(k,r)$$

$$= \Omega\,(k,r_1) - \Omega\,(k,r_0) \ .$$

Nach (11) ist

14)
$$\Omega\,(k,r_1) - \Omega\,(k,r_0) \leq \Omega_b(k,r_1) - \Omega_a(k,r_0) \ .$$

Nun gilt wegen (8)

$$1 - \Omega_b(k,r_1) = 1 - \sum_{r=0}^{r_1} w_b(k,r)$$

$$\leq 1 - \sum_{k(b-q) \leq r \leq k(b+q)} w_b(k,r)$$

$$\leq 1 - \left(1 - \frac{1}{k} \cdot \frac{b(1-b)}{q^2}\right) \quad \text{oder}$$

(15)
$$\Omega_b(k,r_1) \leq 1 - \frac{1}{k} \cdot \frac{b(1-b)}{q^2}$$

Ferner ist wegen (8) :

$$\Omega_a(k,r_0) = \sum_{r=0}^{r_0} w_a(k,r)$$

$$= 1 - \sum_{r=r_0+1}^{k} w_a(k,r)$$

$$\Omega_a(k,r_0) \leq 1 - \sum_{r=r_0+1}^{r_1} w_a(k,r)$$

$$\leq 1 - \sum_{k(a-q)\leq r\leq k(a+q)} w(k,r)$$

$$\leq 1 - \left(1 - \frac{1}{k} \cdot \frac{a(1-a)}{q^2}\right) \quad \text{oder}$$

$$\Omega_a(k,r_0) \leq \frac{1}{k} \cdot \frac{a(1-a)}{q^2} \tag{16}$$

Setzt man (14), (15) und (16) in (13) ein, so ergibt sich :

$$1 \geq \sum_{k(a-q)\leq r\leq k(b+q)} w(k,r) \geq 1 - \frac{1}{k} \cdot \frac{a(1-a) + b(1-b)}{q^2} \tag{17}$$

Aus (17) folgt die Behauptung für festes q und k

Ist nun $a = 0$, $b < 1$, so wählt man q in $0 < q < 1-b$. Aus (13) wird dann

$$\sum_{0\leq r\leq k(b+q)} w(k,r) = \sum_{r=0}^{r_1} w(k,r) \geq \Omega_b(k,r_1)$$

und daraus wegen (15)

$$\sum_{0\leq r\leq k(b+q)} w(k,r) \geq 1 - \frac{1}{k} \cdot \frac{b(1-b)}{q^2}$$

$$= 1 - \frac{1}{k} \cdot \frac{a(1-a) + b(1-b)}{q^2} \quad .$$

Entsprechend zeigt man die Richtigkeit von (17) für $0 < a$, $b = 1$, indem man q in $0 < q < a$ wählt.

§ 8. Ueber die 2. Bernoullische Aussage für ein Verteilungssystem.

Wir schreiben die das Verteilungssystem charakterisierenden Formeln nocheinmal an :

(I) $w(0,0) = 1$

(II) $w(t,r) = (1-U(t,r)) w(t-1,r) + U(t,r-1) w(t-1,r-1)$

$w(t,-1) = w(t,t+1) = 0$

(III) $a \leqq U(t,r) \leqq b$

a und b sind fest; ferner soll jetzt gelten :

(IV) $0 < a \; ; \; b < 1$

Die 2. Bernoullische Aussage würde besagen, dass $\lim_{k=\infty} w(k,r) = 0$ ist, wie auch w(k,r) durch (I) bis (IV) bestimmt wird. Dass für ein allgemeines Verteilungssystem die 2. Bernoullische Aussage nicht gilt, werden wir später an dem Spezialfall $a+b = 1$, $a < b$ zeigen. Vorher soll das Zutreffen der 2. Bernoullischen Aussage für den Poisson'schen Spezialfall bewiesen werden.

In § 6 wurde für den Poisson'schen Spezialfall als erzeugende Funktion

(1) $E(x) = \prod_{s=1}^{t} (1-U(s)+U(s) x)$ eingeführt. Wir setzen

darin nun

$1-U(s) = L_s$; $U(s) = R_s$, sodass also $L_s + R_s = 1$

ist. Formel (II) lautet dann für t+1 :

(2) $$w(t+1,r) = L_{t+1}\, w(t,r) + R_{t+1}\, w(t,r-1)$$

und Formel (1) :

(3) $$E(x) = \prod_{s=1}^{t} (L_s + R_s x)$$

Wir stellen nun folgende Ueberlegungen an :

1. Bei festem k steigt $w(k,r)$ mit wachsendem r bis zu einem gewissen Höchstwert $w_k^* = w(k,r_k)$ an und nimmt von da an dauernd ab, d.h. also

(4) $$w(k,0) \leq w(k,1) \leq \ldots \leq w(k,r_k)$$
$$w(k,r_k) \geq \ldots \geq w(k,k-1) \geq w(k,k)$$

Den Beweis von (4) führen wir durch vollständige Induktion :

Für $k = 1$ ist (4) trivial; (4) sei für $t \leq k$ bereits bewiesen. Dann wird

$$\frac{w(k+1,r)}{w(k+1,r-1)} = \frac{L_{k+1}\, w(k,r) + R_{k+1}\, w(k,r-1)}{L_{k+1}\, w(k,r-1) + R_{k+1}\, w(k,r-2)}$$

Ist $r \leqq r_k$, so gilt $w(k+1,r) \geqq w(k+1,r-1)$.

Ist $r-2 \geqq r_k$, so gilt $w(k+1,r) \leqq w(k+1,r-1)$.

Es bleibt also lediglich noch der Fall $r_k = r-1$ übrig.

Es können drei Fälle eintreten :

$w(k+1,r_k+1) \gtreqless w(k+1,r_k)$

Gilt das Zeichen $>$ oder das Zeichen $=$, so ist

$r_{k+1} = r_k+1$; gilt das Zeichen $<$, so ist

$r_{k+1} = r_k$. Jedenfalls gibt es stets ein r_{k+1} ,

sodass (4) auch für k+1 und damit allgemein gilt .

Die Beziehung $r_{k+1} = \begin{cases} r_k \\ r_k+1 \end{cases}$ drücken wir so aus,

dass die Folge der Maximastellen r_k im Galton'schen Brett eine Kette bilde. Wie dies zu verstehen ist, wird durch die Fig. 1 veranschaulicht:

(Fig.1)

Zu jedem Glied der Kette gibt es ein Glied in der vorhergehenden Reihe und dieses liefert einen Beitrag zu dem betrachteten Glied.

2. Die Werte w_k^* bilden eine monoton fallende Reihe; wir unterscheiden dabei zwei Fälle :

a) $r_{k+1} = r_k$

$$w_{k+1}^* = L_{k+1}\, w_k^* + R_{k+1}\, w(k, r_k - 1)$$

$$= w_k^* - R_{k+1}\left[w_k^* - w(k, r_k - 1)\right] \leq w_k^*$$

b) $r_{k+1} = r_k + 1$

$$w_{k+1}^* = L_{k+1}\, w(k, r_k + 1) + R_{k+1}\, w_k^*$$

$$= w_k^* - L_{k+1}\left[w_k^* - w(k, r_k + 1)\right] \leq w_k^*$$

Eine monoton fallende Zahlenfolge konvergiert :

(5) $$\lim_{k=\infty} w_k^* = \sigma$$

3. Wir führen eine Folge

$$r_k^{(1)} = \begin{cases} r_k - 1 & \text{, wenn } r_{k+1} = r_k \text{ ,} \\ r_k + 1 & \text{, wenn } r_{k+1} = r_k + 1 \text{ ist,} \end{cases}$$

ein. Wir unterscheiden wieder :

a) $r_{k+1} = r_k$

Aus (2) folgt

$$w(k+1, r_{k+1}) = L_{k+1}\, w(k, r_k) + R_{k+1}\, w(k, r_k - 1)$$

$$w^*_{k+1} = L_{k+1}\, w^*_k + R_{k+1}\, w(k, r_k^{(1)})$$

$$= w^*_k - R_{k+1}\left[w^*_k - w(k, r_k^{(1)})\right]$$

und hieraus wegen (III) und (IV)

$$w^*_k - w(k, r_k^{(1)}) = \frac{1}{R_{k+1}}\,(w^*_k - w^*_{k+1}) \tag{6'}$$

b) $r_{k+1} = r_k + 1$

Aus (2) folgt

$$w(k+1, r_{k+1}) = L_{k+1}\, w(k, r_k + 1) + R_{k+1}\, w(k, r_k)$$

$$w^*_{k+1} = L_{k+1}\, w(k, r_k^{(1)}) + R_{k+1}\, w^*_k$$

$$= w^*_k - L_{k+1}\left[w^*_k - w(k, r_k^{(1)})\right]$$

und hieraus wegen (III) und (IV)

$$w^*_k - w(k, r_k^{(1)}) = \frac{1}{L_{k+1}}\,(w^*_k - w^*_{k+1}) \tag{6''}$$

Aus (5) folgt, dass es zu jedem vorgegebenen ε' ein K derart gibt, dass $w^*_k - w^*_{k+1} < \varepsilon'$ für alle $k \geq K$ gilt.

Ist nun $\epsilon > 0$ gegeben, so wird ϵ' so klein gewählt, dass sowohl $\frac{\epsilon'}{a}$ als auch $\frac{\epsilon'}{1-b}$ kleiner als ϵ ist. Dann gilt für alle $k \geq K$

$$\frac{\epsilon'}{R_{k+1}} < \frac{\epsilon'}{a} \quad , \quad \frac{\epsilon'}{L_{k+1}} < \frac{\epsilon'}{1-b} \quad , \text{ also}$$

$$\left| w_k^* - w(k, r_k^{(1)}) \right| < \epsilon \qquad \text{oder}$$

$$\lim_{k=\infty} w_k^* = \lim_{k=\infty} w(k, r_k^{(1)}) = \sigma \quad .$$

4. In Fig. 2 sind die Punkte r_k durch Kreuz und $r_k^{(1)}$ durch Kreis bezeichnet. Zu dieser Doppelkette konstruiert man nun eine Folge $r_k^{(2)}$ in der in der Figur angedeuteten Weise; dabei ist Bedingung, dass Beiträge zu den Punkten r_{k+1} und $r_{k+1}^{(1)}$ nur von den Punkten r_k, $r_k^{(1)}$ und $r_k^{(2)}$ zufliessen sollen.

(Fig.2)

Entsprechend wie in Nr. 3 zeigt man mit Hilfe von (II) bezw. (2) und dann von (III) und (IV), dass

$$\lim_{k=\infty} w(k, r_k^{(2)}) = \lim_{k=\infty} w_k^* = \sigma \quad .$$

Indem man so weitermacht, gelangt man zu einer j-fachen Kette. Für die Werte $w(k, r_k^{(\lambda)})$ gilt dann

$$\lim_{k=\infty} w(k, r_k^{(\lambda)}) = \sigma \quad , \qquad \lambda = 1, 2, \ldots , j \; .$$

Man kann also j+1 nebeneinanderliegende Werte

r_k, $r_k^{(1)}$, ... , $r_k^{(j)}$ in jeder Zeile derart angeben, dass

$\lim_{k=\infty} \sum_{\lambda=0}^{j} w(k, r_k^{(\lambda)}) = (j+1) \sigma$ ist, wobei $r_k = r_k^{(0)}$ gesetzt wird. Wegen $\sum_{r=0}^{k} w(k, r) = 1$ wird

$$(j+1) \sigma \leqq 1, \text{ also } \quad \sigma \leqq \frac{1}{j+1} \quad .$$

Da j beliebig ist, so gilt $\sigma = 0$ oder

$$\lim_{k=\infty} w_k^* = 0 \quad .$$

Dies ist aber der Inhalt der 2. Bernoullischen Aussage für den Poisson'schen Spezialfall.

Wir wollen nun einen Spezialfall betrachten, in dem die 2. Bernoullische Aussage nicht mehr gilt. Dazu konstruieren

wir im Rahmen der Formeln (I) bis (IV) zu vorgegebenem k_0 eine " Maximal-Verteilung " $w^{(k_0)}(k_0,r)$ derart, dass für jedes andere Verteilungssystem w(k,r) mit denselben Konstanten a, b gilt :

$$\operatorname*{Max}_{r=0,1,\ldots,k_0} w^{(k_0)}(k_0,r) \geq w(k_0,\rho)$$

1. Aufgabe : Bei gegebenen t_0 und r_0 mit $0 \leq r_0 \leq t_0$ ist durch geeignete Wahl der U(t,r) im Rahmen der Bedingung (III) $w(t_0,r_0)$ zu einem Maximum zu machen.

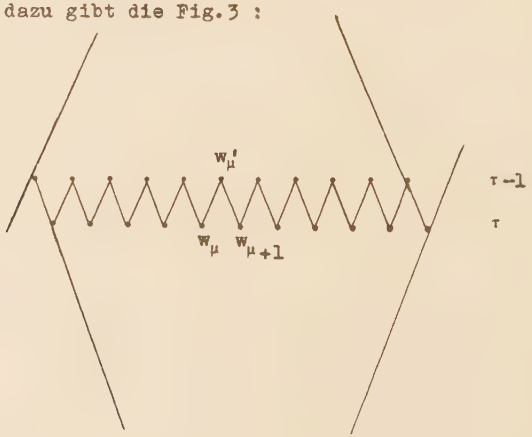
Es ist wieder zweckmässig, die Bezeichnungen L und R einzuführen, jedoch in etwas geänderter Weise :

$1 - U(t,r) = L(t-1,r)$

$U(t,r) = R(t-1,r)$

Für $t' \geq \tau$ seien die Werte L(t', ρ) und R(t', ρ) bereits festgelegt. Nun sollen die Werte L(τ -1, ρ) und R(τ -1, ρ) so bestimmt werden, dass $w(t_0,r_0)$ maximal ausfällt. Die Anleitung dazu gibt die Fig.3 :

(Fig.3)



Die in der Zeile τ stehenden , zu $w(t_o, r_o)$ beitragenden Werte $w(\tau, \rho)$ werden mit $w_1, w_2, \ldots, w_\lambda$ bezeichnet. Man kann dann $w(t_o, r_o)$ als lineare Funktion der Werte $w_1, w_2, \ldots, w_\lambda$ anschreiben :

(7) $$w(t_o, r_o) = C_1 w_1 + C_2 w_2 + \ldots + C_\lambda w_\lambda$$

Die Konstanten $C_1, C_2, \ldots, C_\lambda$ ergeben sich mittels (II) aus den bereits festgelegten Werten $L(t', \rho)$ und $R(t', \rho)$.

Analog seien $w'_1, w'_2, \ldots, w'_{\lambda'}$ mit $\lambda - 1 \leqq \lambda' \leqq \lambda + 1$ diejenigen in der Zeile $\tau - 1$ stehenden Werte $w(\tau - 1, \rho)$, die zu $w(t_o, r_o)$ beitragen.

w'_μ gibt $L_\mu w'_\mu$ nach links unten an w_μ und $R_\mu w'_\mu$ nach rechts unten an $w_{\mu+1}$ ab; gegebenenfalls trifft nur eines zu. In diesem letzteren Fall ist L_μ bezw. R_μ maximal zu wählen. In jedem anderen Fall ist der Beitrag von w'_μ zu $w(t_o, r_o)$ gleich

$$(L_\mu C_\mu + R_\mu C_{\mu+1}) w'_\mu = \left\{ C_\mu + R_\mu (C_{\mu+1} - C_\mu) \right\} w'_\mu \quad .$$

Dieser Beitrag wird am grössten, wenn

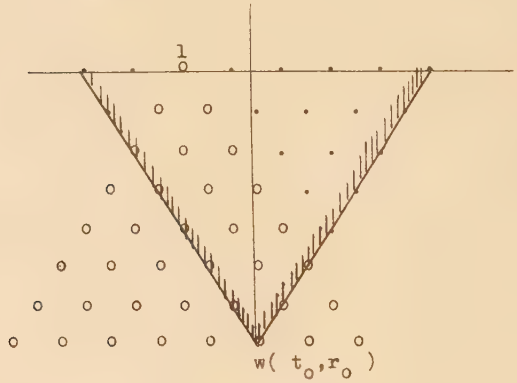
für $C_{\mu+1} - C_\mu > 0$ R_μ maximal, also gleich b

für $C_{\mu+1} - C_\mu < 0$ R_μ minimal, also gleich a

gewählt wird. Ist $C_{\mu+1} - C_\mu = 0$, so ist die Wahl von R_μ ohne Einfluss auf den Beitrag von w'_μ . Indem man so von unten her Zeile für Zeile die L und R bestimmt, erhält man dasjenige System der $U(t, r)$, bei dem $w(t_o, r_o)$ den Maximalwert $\overline{w}(t_o, r_o)$ erreicht.

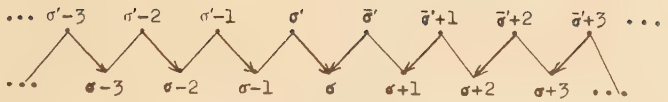
Wir wollen jetzt die angegebene Konstruktion der U(t,r) beim Spezialfall a+b = 1 durchführen. Dabei können wir eine Symmetrieeigenschaft der Konstanten R und L und damit der C beweisen. Es gilt nämlich, dass die C symmetrisch zu einer durch den betrachteten Punkt t_o, r_o gehenden Vertikalen liegen und nach beiden Seiten abfallen. Dazu betrachten wir wieder die Darstellung im Galton'schen Brett.

(Fig.4)



Zu w(t_o,r_o) tragen nur die in dem stark umrandeten Dreieck und auf dem Rand liegenden durch Kreise bezeichneten Punkte bei. Wir erweitern nun das Galton'sche Brett, indem wir in den übrigen Punkten mit $\rho \geq \tau$ (ρ kann grösser als τ sein !) $w(\tau, \rho)$ null setzen. Für gerades τ geben wir ρ in den auf der Vertikallinie durch t_o, r_o liegenden Punkten den Wert σ , für ungerades bezeichnen wir den links von der Vertikallinie liegenden Punkt mit σ und den Punkt weiter rechts mit $\bar{\sigma} = \sigma + 1$. Wir greifen nun zwei Zeilen aus dem umrahmten Dreieck heraus und unterscheiden den Fall : τ gerade und den Fall : τ ungerade.

(Fig.5)



Ist die behauptete Eigenschaft der C für $t' \geqq \tau$ bereits bewiesen, so ergibt sich die angedeutete Festlegung der R und L. In der durch einen Pfeil angedeuteten Richtung wird der Bruchteil b und damit, wegen a+b = 1 , in der anderen Richtung der Bruchteil a abgegeben. Es gilt dann, wenn der Akzent wieder die Zeile τ -1 andeutet:

(8) $$C'_{\sigma-\rho} = a\, C_{\sigma-\rho-1} + b\, C_{\sigma-\rho} \quad , \qquad \rho > 0$$

(9) $$C'_{\sigma} = a\, C_{\sigma-1} + b\, C_{\sigma}$$

(10) $$C'_{\bar{\sigma}} = a\, C_{\sigma+1} + b\, C_{\sigma}$$

(11) $$C'_{\bar{\sigma}+\rho} = a\, C_{\sigma+\rho+1} + b\, C_{\bar{\sigma}+\rho} \quad , \qquad \rho > 0$$

Aus diesen Beziehungen ergibt sich die Behauptung ohne weiteres. Damit ist der Beweis für den Fall : τ gerade geführt. Im Fall : τ ungerade lassen sich ohne weiteres entsprechende Schlüsse ziehen.

2. Aufgabe : Von den nach der 1. Aufgabe bestimmten maximalen $\bar{w}(t_o, r_o)$ soll bei festem t_o und veränderlichem r_o das grösste bestimmt werden.

Nach der 1. Aufgabe können wir bei der durch die starken Linien in Fig. 4 angedeuteten Erweiterung des Galton'schen Bretts die Beiträge $C_\rho^{(t)}$ bestimmen, die die Werte w auf der Grundlinie (t = 1) zu $\overline{w}(t_o,r_o)$ leisten. Es ist

$$\overline{w}(t_o,r_o) = C_o^{(t_o)} w_o + C_1^{(t_o)} w_1 + \ldots + C_{t_o}^{(t_o)} w_{t_o}$$

Die Werte w sind nun sämtlich 0 mit der einzigen Ausnahme $w(1,1) = 1$. Da die Verbindungslinie der Punkte mit den Werten $w(1,1)$ und $w(t,t)$ parallel mit der Verbindungslinie der Punkte mit den Werten w_o und $w(t_o,r_o)$ läuft, so ist der Abstand der Punkte w_o und $w(1,1)$ gleich dem Abstand von $w(t_o,r_o)$ und $w(t_o,t_o)$, also gleich $t_o - r_o$. Es ist also $w(1,1) = w_{t_o-r_o} = 1$ und damit

$$\overline{w}(t_o,r_o) = C_{t_o-r_o}^{(t_o)} .$$

Im Fall a+b = 1 ist nach der 1. Aufgabe :

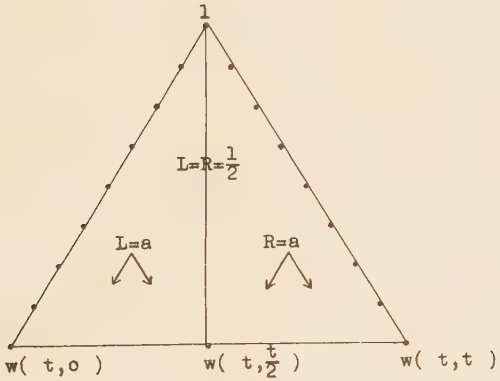
$$\operatorname*{Max}_{r=0,1,\ldots,t_o} C_r^{(t_o)} = C_{\left[\frac{t}{2}\right]}^{(t_o)}$$

$$\operatorname*{Max}_{r=0,1,\ldots,t_o} \overline{w}(t_o,r) = C_{\left[\frac{t}{2}\right]}^{(t_o)}$$

Wir können darnach die gewünschte Maximalverteilung für a+b = 1 angeben: Durch Fig. 6a und Fig. 6b lässt sich $w^{(k_o)}(k,r)$ im Galton'schen Brett darstellen.

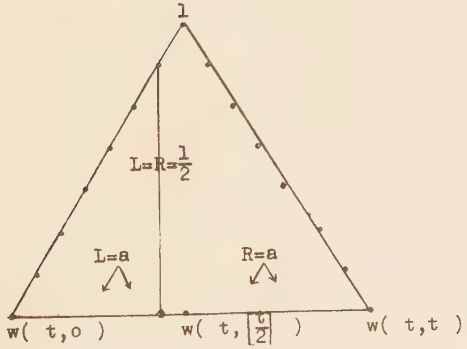
(Fig.6a)

k_o gerade .



(Fig.6b)

k_o ungerade .

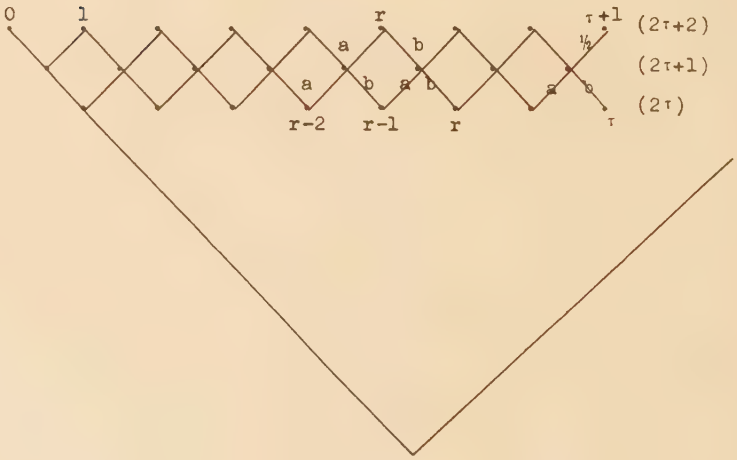


Nunmehr sollen die Grössen $C_r^{(t)}$ berechnet werden und zwar getrennt für gerade und für ungerade t . Dazu leiten wir zunächst drei Rekursionsformeln ab, mit deren Hilfe dann die Formeln für $C_r^{(t)}$ berechnet werden. Symmetriehalber beschränken wir uns auf $0 \leqq r \leqq \left[\frac{t}{2}\right]$. Es ist

$w(0,0) = C_0^{(0)} = 1$.

Für die Ableitung der Rekursionsformeln benützen wir die Figuren 4 und 5 und die daran geknüpften Ueberlegungen ; die Formeln (8) bis (11) gelten auch für die $C_r^{(t)}$. Aus Fig.7 lassen sich die Zusammenhänge ablesen :

(Fig.7)



Aus der Figur liest man ab :

(12) $$c_r^{(2\tau+1)} = a\, c_{r-1}^{(2\tau)} + b\, c_r^{(2\tau)} \quad ; \quad 1 \leq r \leq \tau$$

$$c_0^{(2\tau+1)} = b\, c_0^{(2\tau)}$$

(13) $$c_r^{(2\tau+2)} = a^2\, c_{r-2}^{(2\tau)} + 2ab\, c_{r-1}^{(2\tau)} + b^2\, c_r^{(2\tau)} \quad ; \quad 2 \leq r \leq \tau$$

$$c_1^{(2\tau+2)} = 2ab\, c_0^{(2\tau)} + b^2\, c_1^{(2\tau)}$$

$$c_0^{(2\tau+2)} = b^2\, c_0^{(2\tau)}$$

(14) $$c_{\tau+1}^{(2\tau+2)} = a\, c_{\tau-1}^{(2\tau)} + b\, c_\tau^{(2\tau)}$$

Mit Hilfe dieser Formeln lassen sich nun die Werte $c_r^{(t)}$ berechnen :

$$c_0^{(0)} = 1$$

$$c_0^{(1)} = b$$

$$c_0^{(2)} = b^2 \quad ; \quad c_1^{(2)} = b$$

$$c_0^{(3)} = b^3 \quad ; \quad c_1^{(3)} = ab^2 + b^2$$

$$c_0^{(4)} = b^4 \quad ; \quad c_1^{(4)} = 2ab^3 + b^3 \quad ; \quad c_2^{(4)} = ab^2 + b^2$$

.....

Ehe wir nun die allgemeinen Formeln für die $c_r^{(t)}$ aufstellen, führen wir noch eine Hilfsgrösse $k_{t,r}$ durch folgende

Gleichungen ein :

$$k_{t,r} = \binom{t}{r} - 2\binom{t-1}{r-1} = (1 - \frac{2r}{t})\binom{t}{r} \text{ für } 0 \leq r \leq \left[\frac{t}{2}\right] ; \quad t \geq 1$$

$$k_{t,r} = 0 \quad \text{für} \quad r < 0 \quad .$$

Da $\binom{t}{r} + \binom{t}{r+1} = \binom{t+1}{r+1}$ ist, so gilt

(15) $$k_{t,r} + k_{t,r-1} = k_{t+1,r} \quad .$$

Durch zweimalige Anwendung von (15) folgt :

(15') $$k_{t,r} + 2\,k_{t,r-1} + k_{t,r-2} = k_{t+2,r}$$

Ferner ist für $s \geq 2$

$$k_{2s-1,s-2} + k_{2s-2,s-2} = \frac{3}{2s-1}\binom{2s-1}{s-2} + \frac{2}{2s-2}\binom{2s-2}{s-2}$$

$$= \frac{(2s-2)!}{(s-2)!\,s!}\left(\frac{3}{s+1} + \frac{1}{s-1}\right)$$

$$= \frac{(2s-1)!}{(s-1)!\,(s+1)!}\,\frac{2}{2s}\,2s$$

$$= (1 - \frac{2(s-1)}{2s})\binom{2s}{s-1} = k_{2s,s-1}$$

(16) $$k_{2s-1,s-2} + k_{2s-2,s-2} = k_{2s,s-1}$$

Diese Formel wird nun für $s = \tau + 1$ angeschrieben :

$$k_{2\tau+1,\tau-1} + k_{2\tau,\tau-1} = k_{2\tau+2,\tau}$$

Wegen (15) ist :

$$k_{2\tau+1,\tau-1} + k_{2\tau,\tau-1} = k_{2\tau,\tau-1} + k_{2\tau,\tau-2} + k_{2\tau,\tau-1} = k_{2\tau+2,\tau}$$

oder

(17) $$k_{2\tau,\tau-2} + 2\,k_{2\tau,\tau-1} = k_{2\tau+2,\tau}$$

Schliesslich benötigt man noch

18) $k_{2\tau-2,\tau-2} = k_{2\tau-1,\tau-1}$

Nach (16) ist nämlich für $s = \tau$:

$k_{2\tau-1,\tau-2} + k_{2\tau-2,\tau-2} = k_{2\tau,\tau-1}$

und wegen (15) :

$k_{2\tau,\tau-1} = k_{2\tau-1,\tau-1} + k_{2\tau-1,\tau-2}$

Aus den beiden Gleichungen folgt (18) .

Tabelle der $k_{t,r}$:

	r=0	r=1	r=2	r=3	r=4
t= 1	1				
2	1				
3	1	1			
4	1	2			
5	1	3	2		
6	1	4	5		
7	1	5	9	5	
8	1	6	14	14	
9	1	7	20	28	14
10	1	8	27	48	42

Mit Hilfe der Koeffizienten $k_{t,r}$ lässt sich nun der Wert $C_r^{(t)}$ angeben. Für $\tau \geq 2$ ist :

$$C_\tau^{(2\tau)} = b^\tau \left\{ \sum_{\nu=0}^{\tau-2} k_{\tau+\nu,\nu}\, a^\nu + k_{2\tau-2,\tau-2}\, a^{\tau-1} \right\} \tag{19}$$

$$C_r^{(2\tau)} = b^{2\tau-r} \left\{ \sum_{\nu=0}^{r} k_{2\tau-r+\nu,\nu}\, a^\nu \right\} , \quad 0 \leq r < \tau . \tag{20}$$

Beweis :

Im Falle $\tau = 2$ treffen die Formeln (19) und (20) zu. Es sei nun $\tau \geq 2$ und ihre Richtigkeit für $\tau' \leq \tau$ bereits gezeigt. Aus (14) folgt dann :

$$C_{\tau+1}^{(2\tau+2)} = a\, b^{\tau+1} \left\{ \sum_{\nu=0}^{\tau-1} k_{\tau+\nu+1,\nu}\, a^\nu \right\}$$

$$+ b^{\tau+1} \left\{ \sum_{\nu=0}^{\tau-2} k_{\tau+\nu,\nu}\, a^\nu + k_{2\tau-2,\tau-2}\, a^{\tau-1} \right\}$$

$$= b^{\tau+1} \left\{ \sum_{\nu=0}^{\tau} k_{\tau+\nu,\nu-1}\, a^\nu + \sum_{\nu=0}^{\tau-2} k_{\tau+\nu,\nu}\, a^\nu \right.$$

$$\left. + k_{2\tau-2,\tau-2}\, a^{\tau-1} \right\}$$

$$= b^{\tau+1} \left\{ \sum_{\nu=0}^{\tau-2} k_{\tau+1+\nu,\nu}\, a^\nu + \left(k_{2\tau-1,\tau-2} + k_{2\tau-2,\tau-2} \right) a^{\tau-1} \right.$$

$$\left. + k_{2\tau,\tau-1}\, a^\tau \right\}$$

$$C_{\tau+1}^{(2\tau+2)} = b^{\tau+1} \left\{ \sum_{\nu=0}^{\tau+1-2} k_{\tau+1+\nu,\nu}\, a^{\nu} + k_{2(\tau+1)-2,\tau+1-2}\, a^{\tau+1-1} \right\}$$

Entsprechend zeigt man mit Hilfe von (13) für $2 \leq r < \tau$

$$C_{r}^{(2\tau+2)} = a^2\, b^{2\tau-r+2} \left\{ \sum_{\nu=0}^{r-2} k_{2\tau-r+2+\nu,\nu}\, a^{\nu} \right\}$$

$$+ 2ab\, b^{2\tau-r+1} \left\{ \sum_{\nu=0}^{r-1} k_{2\tau-r+1+\nu,\nu}\, a^{\nu} \right\}$$

$$+ b^2\, b^{2\tau-r} \left\{ \sum_{\nu=0}^{r} k_{2\tau-r+\nu,\nu}\, a^{\nu} \right\}$$

$$= b^{2\tau+2-r} \left\{ \sum_{\nu=0}^{r} k_{2\tau-r+\nu,\nu-2}\, a^{\nu} \right\}$$

$$+ b^{2\tau+2-r} \left\{ \sum_{\nu=0}^{r} 2\, k_{2\tau-r+\nu,\nu-1}\, a^{\nu} \right\}$$

$$+ b^{2\tau+2-r} \left\{ \sum_{\nu=0}^{r} k_{2\tau-r+\nu,\nu}\, a^{\nu} \right\}$$

$$= b^{2(\tau+1)-r} \left\{ \sum_{\nu=0}^{r} k_{2(\tau+1)-r+\nu,\nu}\, a^{\nu} \right\}$$

Im Falle $\tau = r$ geht der Beweis ganz analog; man verwendet die Formeln (17) und (18).

Nunmehr soll $\lim C_\tau^{(2\tau)}$ bestimmt werden. Dazu benötigen wir den Wert der Differenz $C_\rho^{(2\rho)} - C_{\rho-1}^{(2\rho)}$. Aus (19) und (20) folgt :

$$C_\rho^{(2\rho)} - C_{\rho-1}^{(2\rho)} = b^\rho \left\{ \sum_{\nu=0}^{\rho-2} k_{\rho+\nu,\nu}\, a^\nu + k_{2\rho-2,\rho-2}\, a^{\rho-1} \right\}$$

$$- b^{\rho+1} \left\{ \sum_{\nu=0}^{\rho-1} k_{\rho+\nu+1,\nu}\, a^\nu \right\}$$

$$= b^\rho \left\{ \sum_{\nu=0}^{\rho-2} k_{\rho+\nu,\nu}\, a^\nu + k_{2\rho-2,\rho-2}\, a^{\rho-1} \right.$$

$$\left. + \sum_{\nu=0}^{\rho-1} k_{\rho+\nu+1,\nu}\, a^{\nu+1} - \sum_{\nu=0}^{\rho-1} k_{\rho+\nu+1,\nu}\, a^\nu \right\}$$

$$= b^\rho \left\{ \sum_{\nu=0}^{\rho-2} \left(k_{\rho+\nu,\nu} + k_{\rho+\nu,\nu-1} - k_{\rho+\nu+1,\nu} \right) a^\nu \right.$$

$$+ \left(k_{2\rho-1,\rho-2} + k_{2\rho-2,\rho-2} \right) a^{\rho-1}$$

$$\left. + k_{2\rho,\rho-1}\, a^\rho - k_{2\rho,\rho-1}\, a^{\rho-1} \right\}$$

$$= k_{2\rho,\rho-1}\, a^\rho\, b^\rho \quad ,$$

wobei $-b = a-1$ und nach (15) bezw. (16)

$k_{\rho+\nu,\nu} + k_{\rho+\nu,\nu-1} - k_{\rho+\nu+1,\nu} = 0$ und

$k_{2\rho-1,\rho-2} + k_{2\rho-2,\rho-2} - k_{2\rho,\rho-1} = 0$ benützt wurden.

Nun gehen wir von der Formel (14) aus :

(21) $$c_\rho^{(2\rho)} = c_{\rho-1}^{(2\rho-2)} - a\,\left(c_{\rho-1}^{(2\rho-2)} - c_{\rho-2}^{(2\rho-2)} \right)$$

Schreibt man (21) der Reihe nach für $\rho = 2, 3, \ldots, \tau$ an, so erhält man durch Addition der linken und rechten Seiten :

$$c_\tau^{(2\tau)} = c_1^{(2)} - a \sum_{\rho=1}^{\tau-1} \left\{ c_\rho^{(2\rho)} - c_{\rho-1}^{(2\rho)} \right\} \quad , \text{ oder}$$

(22) $$c_\tau^{(2\tau)} = b - a \sum_{\rho=1}^{\tau-1} k_{2\rho,\rho-1}\, a^\rho\, b^\rho \quad .$$

Da in (21) $c_{\tau-1}^{(2\tau-2)} - c_{\tau-2}^{(2\tau-2)} > 0$ ist, so bilden die $c_\tau^{(2\tau)}$ eine monoton fallende Folge positiver Zahlen, streben also gegen einen Grenzwert Γ . Man kann also in (22) zur Grenze übergehen :

$$\Gamma = b - a \sum_{\rho=1}^{\infty} k_{2\rho,\rho-1}\, a^\rho\, b^\rho$$

Nun ist $\binom{1/2}{\rho+1} = \frac{1}{2} \left(-\frac{1}{4}\right)^\rho \frac{(2\rho)!}{\rho!\,(\rho+1)!}$, also, da

$k_{2\rho,\rho-1} = \binom{2\rho}{\rho-1} \frac{2}{2\rho} = \frac{(2\rho)!}{\rho!\,(\rho+1)!}$, nach (22)

$$c_\tau^{(2\tau)} = b - 2a \sum_{\rho=1}^{\tau-1} \binom{1/2}{\rho+1} (-4ab)^\rho$$

$$= b + \frac{1}{2b} \sum_{\rho=2}^{\tau} \binom{1/2}{\rho} (-4ab)^\rho \quad .$$

Andererseits ist, da für $a < b$

$0 < 4ab = 4b\,(1-b) = 1 - (2b-1)^2 < 1$,

$$2b-1 = \sqrt{1 - 4ab} = \sum_{\rho=0}^{\infty} \binom{1/2}{\rho} (-4ab)^{\rho} \quad .$$

und die Reihe konvergent. Daraus folgt :

$$r = \lim_{\tau=\infty} C_{\tau}^{(2\tau)} = b + \frac{1}{2b}\left[(2b-1) - 1 + 2ab\right]$$

$$= b + \frac{2b\,(1+a) - 2}{2b} = b + \frac{2\,(1-a^2) - 2}{2b}$$

$$= \frac{b^2 - a^2}{b} = \frac{b - a}{b} \quad .$$

Da die Gleichung (14) für jedes τ gilt, so gilt sie auch in der Grenze :

$$\lim_{\tau=\infty} C_{\tau-1}^{(2\tau)} = \frac{1}{a}\,\frac{b-a}{b} - \frac{b}{a}\,\frac{b-a}{b} = \frac{b-a}{b}\,\frac{1-b}{a} = \frac{b-a}{b}$$

Ebenso lässt sich der Grenzübergang in (13) durchführen für $r = \tau$:

$$\lim_{\tau=\infty} C_{\tau-2}^{(2\tau)} = \frac{1}{a^2}\,\frac{b-a}{b} - \frac{2ab}{a^2}\,\frac{b-a}{b} - \frac{b^2}{a^2}\,\frac{b-a}{b}$$

$$= \frac{b-a}{b\,a^2}\,(\,1 - 2ab - b^2\,) = \frac{b-a}{b}\,\frac{a^2}{a^2} = \frac{b-a}{b}$$

In dieser Weise kann man weitermachen :

Für jedes feste λ_0 gilt

$$\lim_{\tau=\infty} C_{\tau-\lambda_0}^{(2\tau)} = \lim_{\tau=\infty} C_{\tau}^{(2\tau)} = \frac{b-a}{b} \quad .$$

Schliesslich folgt aus(12) für festes λ_0 :

$$\lim_{\tau=\infty} C_{\tau-\lambda_0+1}^{(2\tau+1)} = a \lim_{\tau=\infty} C_{\tau-\lambda_0}^{(2\tau)} + b \lim_{\tau=\infty} C_{\tau-\lambda_0+1}^{(2\tau)}$$

$$= \lim_{\tau=\infty} C_{\tau-\lambda_0}^{(2\tau)} = \frac{b-a}{b}$$

Das Maximum $\frac{b-a}{b}$ kann also auch seitlich in endlich bleibendem Abstand erreicht werden.

Die 2. Bernoullische Aussage gilt also schon im Spezialfall $a+b = 1$, $a < b$ nicht, daher sicher nicht allgemein.

§ 9. Die Bernoullischen Aussagen für X-Folgen.

Die das Verteilungssystem charakterisierenden Formeln (I) bis (III) gelten für jede in § 6 beschriebene Teilfolgenauswahl. Um die Ergebnisse der §§ 7 und 8 auf X-Folgen zu übertragen, treffen wir nun spezielle Teilfolgenauswahlen. Dazu gehen wir von einer X-Folge $E : E_1, E_2, E_3, \ldots$ aus und betrachten das zu E gehörende $P^{(k)}$-Feld, wobei zunächst k fest sein soll:

$E_1 , E_2 , E_3 , \ldots$
$E_2 , E_3 , E_4 , \ldots$
.....
$E_k , E_{k+1} , E_{k+2} , \ldots$

Das Verhältnis der Anzahl der Merkmale m in der n-ten Spalte zu k sei $h_{n,k}$. Es sei $0 < a \leq b < 1$; ferner sei q mit $0 < a-q < b+q < 1$ eine feste Zahl. Dabei sei wieder $a = g^*$, $b = G^*$ gesetzt. Wir betrachten nun das Intervall $\left[a-q\ ,\ b+q\right]$ und stellen für jedes n fest, ob $h_{n,k}$ in diesem Intervall liegt oder nicht. Jetzt bilden wir ein neues Alternativ $A^{(k)} : A_1^{(k)}\ ,\ A_2^{(k)}\ ,\ A_3^{(k)}\ ,\ \ldots\ ,$ indem wir

$A_n^{(k)} = 1$ setzen, wenn $a-q \leq h_{n,k} \leq b+q$ ist, und

$A_n^{(k)} = 0$ setzen, wenn $h_{n,k} < a-q$ oder $h_{n,k} > b+q$ ist .

Diesem Alternativ $A^{(k)}$ kommt nun ein W-Band $B(\ A^{(k)};\ "1"\ )$ zu. Die 1. Bernoullische Aussage behauptet, dass bei festem q gilt :

(1) $$\lim_{k=\infty} B(\ A^{(k)};\ "1"\ ) = 1$$

Anders ausgedrückt :

Das W-Band $B(\ A^{(k)};\ "1"\ )$, das die Grenzwerte der relativen Häufigkeiten dafür enthält, dass in den Spalten des zu E gehörenden $F^{(k)}$-Feldes mindestens $k(a-q)$ und höchstens $k(b+q)$ Elemente m vorkommen, strebt mit wachsendem k bei festem q gegen den Punkt 1. In dieser Form geht die 1. Bernoullische Aussage unmittelbar in das Bernoullische Gesetz für Z-Folgen über, wenn die schärferen Voraussetzungen der Definition der Z-Folgen gemacht werden (vgl. § 12) .

Um die Behauptung auf die 1. Bernoullische Aussage für ein Verteilungssystem zurückzuführen, betrachten wir die rela-

tive Häufigkeit des Alternativs $A^{(k)}$. Diese ist die relative Anzahl derjenigen Spalten des Feldes, welche mindestens $k(a-q)$ und höchstens $k(b+q)$ Merkmale m enthalten. Es interessiert uns nun die untere Grenzhäufigkeit der Folge $A^{(k)}$. In den Bezeichnungen des § 6 gilt :

(2) $$\underline{\lim_{n=\infty}} h_n(A^{(k)}; \text{"1"}) = \underline{\lim_{n=\infty}} \left\{ \sum_{k(a-q)\leq r\leq k(b+q)} h_n(E^{(k+1)}, m_r^{(k)}) \right\}$$

Es gibt nun eine Folge $n_1', n_2', n_3', \ldots$ wachsender natürlicher Zahlen derart, dass

(3) $$\underline{\lim_{n=\infty}} \left\{ \sum_{k(a-q)\leq r\leq k(b+q)} h_n(E^{(k+1)}, m_r^{(k)}) \right\}$$

$$= \lim_{\nu=\infty} \left\{ \sum_{k(a-q)\leq r\leq k(b+q)} h_{n_\nu'}(E^{(k+1)}, m_r^{(k)}) \right\}$$

wird. Durch weitere Teilfolgenauswahl lässt sich ferner in der in § 6 geschilderten Weise erreichen, dass

$$\lim_{\nu=\infty} h_{n_\nu}(E^{(k)}, m_r^{(k-1)}) = w(k-1,r) \quad \text{für } r = 0, 1, \ldots, k-1$$

und

$$\lim_{\nu=\infty} a_{n_\nu}^{(k,r)} = U(k,r) \quad \text{für } r = 0, 1, \ldots, k$$

existieren. Für jede der so gebildeten Teilstellenauswahlen gilt bei festem k nach (2) und (3)

$$\underline{\lim_{n=\infty}} h_n(A^{(k)}; \text{"1"}) = \lim_{\nu=\infty} \left| \sum_{k(a-q)\leq r\leq k(b+q)} w(k,r) \right|$$

und damit nach § 7 (17) :

$$1 \geq \underline{\lim_{n=\infty}} h_n(A^{(k)}; \text{"1"}) \geq 1 - \frac{1}{k} \cdot \frac{a(1-a)+b(1-b)}{q^2}$$

Es ist also für festes q :

$$\lim_{k=\infty} \left| \underline{\lim_{n=\infty}} h_n(A^{(k)}; \text{"1"}) \right| = 1 \quad .$$

Die 2. Bernoullische Aussage beschäftigt sich mit dem grössten Wert, den w(k,r) für $k \to \infty$ annehmen kann. Es sei $C > 0$ fest. w sei eine beliebige Zahl in $0 \leq w \leq 1$. Wir betrachten nun die obere Grenzhäufigkeit dafür, dass in einer Spalte von $F^{(k)}$ die Anzahl der m von kw um weniger als C abweicht. Diese Grenzhäufigkeit bezeichnen wir mit $\overline{W}(|r-kw| < C)$. Nach § 6 können wir schreiben

$$\overline{W}(|r-kw| < C) = \overline{\lim_{n=\infty}} \sum_{kw-C\leq r\leq kw+C} h_n(E^{(k+1)}, m_r^{(k)})$$

und erreichen durch fortgesetzte Teilfolgenauswahl, dass folgende Grenzwerte gleichzeitig existieren :

$$\lim_{\nu=\infty} \sum_{kw-C\leq r\leq kw+C} h_{n_\nu}(E^{(k+1)}, m_r^{(k)}) = \overline{\lim_{n=\infty}} \sum_{kw-C\leq r\leq kw+C} h_n(E^{(k+1)}, m_r^{(k)})$$

$$\lim_{\nu=\infty} h_{n_\nu}(E^{(k)}, m_r^{(k-1)}) = w(k-1,r)$$

$$\lim_{\nu=\infty} a_{n_\nu}^{(k,r)} = U(k,r)$$

Da es sich um die obere Grenzhäufigkeit handelt, so interessiert uns insbesondere derjenige Wert, für welchen $\overline{W}$ Maximum wird. Sicher ist

$$\overline{W}(|r-kw| < C) < 2C \max_{r=0,1,\dots,k} \left\{ \overline{\lim_{n}} h_n(E^{(k+1)}, m_r^{(k)}) \right\} .$$

In § 8 haben wir für den Spezialfall $a+b=1$ eine solche Maximalverteilung bestimmt und als Wert der grössten oberen Grenzhäufigkeit $C^{(t)}_{[\frac{t}{2}]}$ gefunden. Es ergab sich dabei, dass

$$\lim_{k=\infty} C^{(t)}_{[\frac{t}{2}]} = \frac{b-a}{b}$$

war. Demnach gilt für $a+b=1$:

4) $$\overline{\lim_{k=\infty}} \overline{W} (|r-kw| < C) \leq 2C \frac{b-a}{b}$$

Ob dieses lediglich durch Betrachtung allgemeiner Verteilungssysteme gewonnene Ergebnis sich noch verschärfen lässt, wenn man auf die Herkunft der Verteilungssysteme aus den $F^{(k)}$-Feldern einer X-Folge weitergehend Bezug nimmt, bleibt hier unbeantwortet. Ein feineres Eingehen auf die Entstehung der Verteilungssysteme aus X-Folgen dürfte jedoch grösste Schwierigkeiten bieten. Als Ersatz für die 2. Bernoullische Aussage ist (4) jedenfalls unzulänglich; dies kannals ein Grund dafür angesehen werden, dass in der Wahrscheinlichkeitsrechnung durchweg Z-Folgen und nicht allgemeinere X-Folgen Gegenstand der Betrachtung sind.

§ 10. Unempfindlichkeit und Existenz von X - Folgen.

Wir wollen zeigen, dass das Intervall $[a,b]$ nicht geändert wird, wenn man die Elemente E_{ν_1} , E_{ν_2} , E_{ν_3} , ... , die in der X-Folge E mit der Grenzhäufigkeit 0 vorkommen, streicht oder abändert.

k werde beliebig gewählt. Man betrachtet das Feld $F^{(k)}$

E_1 , E_2 , E_3 , ...

E_2 , E_3 , E_4 , ...

.....

E_k , E_{k+1} , E_{k+2} , ...

Streicht man in E = E' das Element E_ν, so wird damit in E" das Element an der $(\nu-1)$ten Stelle, ... , in $E^{(k)}$ das Element an der $(\nu-k+1)$ten Stelle gestrichen. Nach der Streichung stehen an ν-ter Stelle

in E' das Element $E_{\nu+1}$

in E" das Element $E_{\nu+2}$

.....

in $E^{(k)}$ das Element $E_{\nu+k}$.

Das bedeutet aber : Durch die Streichung eines Elementes E_ν werden alle Spalten des Feldes, welche nach der ν-ten Spalte vorkommen, lediglich um eine Stelle nach links verschoben. Von der Streichung betroffen werden nur die k Spalten, in denen das Element E_ν vorkommt. Wird E_ν durch ein anderes Element ersetzt, so tritt eine Aenderung nur in den k Spalten

ein, in denen E_ν auftritt; eine Verschiebung findet nicht statt.

Werden nun Streichungen oder Abänderungen mit der oberen Grenzhäufigkeit τ vorgenommen, so werden dadurch höchstens die Spalten an den Stellen ν_1-k+1 , ... , ν_1 ; ν_2-k+1 , ... , ν_2 ; ν_3-k+1 , ... , ν_3 ; ... betroffen.

Hat nun die Auswahlfolge $E^{(k,r)}$ das W-Band $g^{(k,r)},G^{(k,r)}$ so tritt eine Abänderung oder Streichung höchstens an den Stellen ein, die in $F^{(k+1)}$ in Spalten an den Stellen ν_1-k , ... , ν_1 ; ν_2-k , ... , ν_2 ; ν_3-k , ... , ν_3 ; ... vorkommen. Die hier auftretenden Zahlen bilden eine Folge natürlicher Zahlen, die in der Folge aller natürlichen Zahlen mit der Grenzhäufigkeit

$$(k+1)\,\tau \doteq \overline{\lim_{n=\infty}} \frac{(k+1)n}{\nu_n}$$

vorkommen. Man kann also § 3 Satz 2 anwenden, wobei das W-Band der geänderten Auswahlfolge $E^{*(k,r)}$ mit $\left[\overline{g},\overline{G}\right]$ bezeichnet werden möge :

$$g^{(k,r)} - (k+1)\,\tau \leq \overline{g} \leq \overline{G} \leq \frac{G^{(k,r)} + \tau\,(k+1)}{1 - (k+1)\,\tau}$$

Ist $\tau = 0$, so gilt für jedes k :

$$g^{(k,r)} = \overline{g} \leq \overline{G} = G^{(k,r)}$$

In dieser Form sprechen wir den Unempfindlichkeitssatz aus :

Satz 5. $E : E_1 , E_2 , E_3 , \ldots$ sei eine X-Folge und $\left[g^*,G^*\right]$ das kleinste Intervall, in dem die W-Bänder von E' und allen Auswahlfolgen $E^{(k,r)}$ enthalten sind. Streicht oder ändert man in E die mit der Grenzhäufigkeit 0 auftretenden Elemente $E_{\nu_1} , E_{\nu_2} , E_{\nu_3} , \ldots$ ab, so ändert sich das Intervall $\left[g^*,G^*\right]$ nicht.

Dass es nicht triviale X-Folgen, d.h. X-Folgen, für die $g^* > 0$ und $G^* < 1$ ist, gibt, wird später gezeigt werden. Bei der Einführung der X-Folgen wurde darauf hingewiesen, dass die X-Folgen einen Spezialfall der E-Folgen darstellen. Unterwirft man die X-Folgen, wie dies im nächsten Abschnitt geschieht, gewissen weiteren Einschränkungen, so gelangt man zu dem Spezialfall der Z-Folgen, deren Existenz für jede beliebig vorgegebene Verteilung bewiesen wird. Damit ist dann aber auch die Existenz nicht trivialer X-Folgen gezeigt.

II. Z - Folgen.

§ 11. W - Folgen.

Im I. Abschnitt wurde gezeigt, dass die 2. Bernoullische Aussage für allgemeine Verteilungssysteme nicht zu gelten braucht, sodass die Frage der Gültigkeit dieser Aussage für allgemeine X-Folgen offen bleibt. Ist aber (in den Bezeichnungen des I. Abschnitts) $g^* = G^*$, schrumpft also das W-Band auf einen Punkt zusammen, so gilt die 2. Bernoullische Aussage.

Wir beschäftigen uns zunächst mit der Konstruktion von X-Folgen, für die $g^* = G^*$ ist. Dazu führen wir den Begriff der W-Folge ein :

Def.V. Eine gegebene E-Folge heisst Wahrscheinlichkeitsfolge, abgekürzt W-Folge, wenn der Grenzwert der relativen Häufigkeit

$$w = \lim_{n=\infty} \frac{H_n(E,m)}{n}$$

existiert. w heisst die (mathematische) Wahrscheinlichkeit für das Auftreten von m in der gegebenen E-Folge.

Aus § 3 folgt, dass der Unempfindlichkeitssatz auch für W-Folgen gilt :

Satz 6. Ist $E : E_1 , E_2 , E_3 , \dots$ eine W-Folge mit der Wahrscheinlichkeit w in Bezug auf m und kommen die Elemente $E_{\nu_1} , E_{\nu_2} , E_{\nu_3} , \dots$ in E mit der Grenzhäufigkeit 0 vor, so ist auch die durch Streichung oder Abänderung der Elemente $E_{\nu_1} , E_{\nu_2} , E_{\nu_3} , \dots$ entstehende Folge E^* in Bezug auf m eine W-Folge mit der Wahrscheinlichkeit w.

In Ergänzung der Ausführungen bei Kamke[7] soll zunächst noch ein Additionsgesetz bewiesen werden :

Satz 7. Besitzt die E-Folge E in Bezug auf die abzählbar unendlich vielen Merkmale $m_1 , m_2 , m_3 , \dots$ die echte Verteilung $w_1 , w_2 , w_3 , \dots$, so gilt das allgemeine Additionsgesetz :

$$w(E, m_{\nu_1} + m_{\nu_2} + m_{\nu_3} + \dots) = w(E, m_{\nu_1}) + w(E, m_{\nu_2}) + w(E, m_{\nu_3}) + \dots$$

für jede Folge $1 \leq \nu_1 < \nu_2 < \nu_3 < \dots$

Zum Beweis setzen wir $M = m_{\nu_1} + m_{\nu_2} + m_{\nu_3} + \dots$. Mit $\lambda_1 , \lambda_2 , \lambda_3 , \dots$ werden diejenigen Zahlen der Folge aller natürlichen Zahlen bezeichnet, welche nach Streichung der Zahlen $\nu_1 , \nu_2 , \nu_3 , \dots$ noch stehen bleiben. Dann sei $\overline{M} = m_{\lambda_1} + m_{\lambda_2} + m_{\lambda_3} + \dots$.
Für jedes m gilt :

$$H_n(E, M) + H_n(E, \overline{M}) = n \quad , \quad \text{also}$$

$$h_n(E, M) = 1 - h_n(E, \overline{M})$$

---

7) vgl. E.Kamke, a.a.O.,S.26.

$$\overline{\lim_{n=\infty}} h_n(E,M) \leq 1 - \underline{\lim_{n=\infty}} h_n(E,\overline{M}) \quad , \quad \text{also}$$

$$\overline{\lim_{n=\infty}} h_n(E,M) \leq 1 - (w_{\lambda_1} + w_{\lambda_2} + w_{\lambda_3} + \dots)$$

$$= w_{\nu_1} + w_{\nu_2} + w_{\nu_3} + \dots$$

Ferner ist

$$\underline{\lim_{n=\infty}} h_n(E,M) \geq w_{\nu_1} + w_{\nu_2} + w_{\nu_3} + \dots$$

daraus folgt

$$\lim_{n=\infty} h_n(E,M) = w_{\nu_1} + w_{\nu_2} + w_{\nu_3} + \dots \quad , \quad \text{w.z.b.w.}$$

Nunmehr soll ein Existenzbeweis für W-Folgen mit vorgeschriebener Verteilung gegeben werden; der Beweis ist demjenigen von Kamke[8] nachgebildet, vermeidet aber die dort befindliche Ungenauigkeit.

1. Wir zeigen zunächst, dass es eine nur aus m und $\overline{m}$ aufgebaute E-Folge gibt, für welche stets

   $H_n(E,m) = [nw]$ ist, wenn $0 < w \leq 1$ ist.

   Ist $w = 1$, so ist die Behauptung trivial, da die Folge, deren sämtliche Elemente m sind, die gewünschte Eigenschaft hat.

---

8) vgl. E.Kamke, a.a.O.,S.102/104.

Für $w<1$ wird die erste Stelle mit $\bar{m}$ besetzt. Es ist $H_n(E,m) = 1.w = 0$, wie verlangt. Nun wird angenommen, dass bereits die n ersten Stellen so mit m und $\bar{m}$ besetzt sind, dass $H_n(E,m) = \lfloor nw \rfloor$ ist. Dann soll die (n+1)te Stelle besetzt werden und zwar

mit m, wenn $\lfloor (n+1)w \rfloor = \lfloor nw \rfloor + 1$, und

mit $\bar{m}$, wenn $\lfloor (n+1)w \rfloor = \lfloor nw \rfloor$ ist.

Im ersten Fall ist

$$H_{n+1}(E,m) = H_n(E,m) + 1 = \lfloor nw \rfloor + 1 = \lfloor (n+1)w \rfloor .$$

Im zweiten Fall ist

$$H_{n+1}(E,m) = H_n(E,m) = \lfloor nw \rfloor = \lfloor (n+1)w \rfloor .$$

Die Behauptung gilt also für n+1, wenn sie für n gilt. Da die Richtigkeit für $n = 1$ bewiesen ist, gilt sie allgemein.

2. Nun seien endlich oder abzählbar unendlich viele Zahlen $w_1, w_2, w_3, \ldots$ gegeben, die sämtlich positiv sind und für die $w_1 + w_2 + w_3 + \ldots = 1$ gilt. Wir konstruieren eine Folge E', die nur aus den Merkmalen $\bar{m}, m_1, m_2, m_3, \ldots$ aufgebaut ist. E' soll dabei in Bezug auf m_μ für $\mu = 1, 2, 3, \ldots$ die Wahrscheinlichkeit w_μ und in Bezug auf $\bar{m}$ die Wahrscheinlichkeit 0 haben.

3. Ersetzt man dann nachträglich in E' die mit der Wahrscheinlichkeit 0 auftretenden Glieder $\bar{m}$ etwa durch m_1, so erhält man eine nur aus $m_1, m_2, m_3, \ldots$ aufgebaute E-Folge mit der echten Verteilung $w_1, w_2, w_3, \ldots$.

4. Um eine Folge E' zu erhalten, werden wir schrittweise vorgehen :

1.Schritt : Nach Ziffer 1 gibt es zu jedem $0 < w_1 \leq 1$ eine nur aus m_1 und $\bar{m}$ aufgebaute Folge F' , für welche

$$H_{n_1}(F', m_1) = \left[n_1 \frac{w_1}{w_1+w_2+w_3+\dots} \right] \text{ ist.}^{9)}$$

2.Schritt : Die in F' mit $\bar{m}$ besetzten Stellen bilden eine Teilfolge F" ; die Stellen von F" werden, statt durchweg mit $\bar{m}$, jetzt nach der Vorschrift der Ziffer 1 mit m_2 und $\bar{m}$ so besetzt, dass

$$H_{n_2}(F'', m_2) = \left[n_2 \frac{w_2}{w_2+w_3+w_4+\dots} \right] \text{ ist.}^{9)}$$

.....

kter Schritt : Die in $F^{(k-1)}$ mit $\bar{m}$ besetzten Stellen bilden eine Teilfolge $F^{(k)}$; die Stellen von $F^{(k)}$ werden, statt durchweg mit $\bar{m}$, jetzt so mit m_k und $\bar{m}$ besetzt, dass

$$H_{n_k}(F^{(k)}, m_k) = \left[n_k \frac{w_k}{w_k+w_{k+1}+w_{k+2}+\dots} \right] \text{ ist.}^{9)}$$

---

9) Die Bezeichnungen n_k beziehen sich auf die Zählung innerhalb der Folge $F^{(k)}$, $k = 1, 2, 3, \dots$.

Das Verfahren lässt sich beliebig weit fortsetzen. Dabei wird eine Stelle höchstens mit einem Merkmal besetzt, da bei späteren Schritten des Verfahrens jeweils nur die bis dahin noch mit $\bar{m}$ besetzten Stellen besetzt werden können.

5. Sind nur endlich viele Merkmale m_1 , m_2 , ... , m_p gegeben, so sind nach p Schritten alle Stellen mit einem der Merkmale besetzt, da

$$H_{n_p}(F^{(p)}, m_p) = \left[n_p \frac{w_p}{w_p} \right] = n_p \quad \text{ist .}$$

6. Sind unendlich viele Merkmale m_1 , m_2 , m_3 , ... gegeben, so findet bei dem in Ziffer 4 angegebenen Verfahren eine Besetzung aller Stellen mit den Merkmalen m_1 , m_2 , m_3 ,.. nicht statt. Z.B. bleibt stets die erste Stelle unbesetzt. Für jedes n lässt sich aber angeben, wieviel und welche Stellen mit $\bar{m}$ besetzt sind.

Wir denken die Zahlen n_1 , n_2 , n_3 , ... so, dass bei festgedachtem n unter n_{k+1} die Anzahl derjenigen unter den n ersten Stellen von E' verstanden wird, die weder mit m_1 , noch mit m_2 , ... , noch mit m_k besetzt sind, wobei $n = n_1$ sein soll. Dabei bleiben wir in Uebereinstimmung mit der vorhin eingeführten Bedeutung der n_k . Es gilt

$$n_1 = n$$

$$n_2 = n_1 - \left[n_1 \frac{w_1}{w_1+w_2+w_3+\dots} \right]$$

.....

$$n_{k+1} = n_k - \left[n_k \frac{w_k}{w_k+w_{k+1}+w_{k+2}+\cdots} \right]$$

.....

Daraus folgt, dass

$n_{k+1} = n_k$, wenn $n_k \frac{w_k}{w_k+w_{k+1}+w_{k+2}+\cdots} < 1$, und dass

$n_{k+1} \leq n_k - 1$, wenn $n_k \frac{w_k}{w_k+w_{k+1}+w_{k+2}+\cdots} \geq 1$, ist.

Setzt man bei festem n jetzt $q_k = n_k \frac{w_k}{w_k+w_{k+1}+w_{k+2}+\cdots}$,

so sind höchstens n Glieder der Folge q_1 , q_2 , q_3 , ... grösser oder gleich 1. Es gibt daher ein k_o derart, dass für alle $k \geq k_o$

$q_k < 1$, aber $q_{k_o-1} \geq 1$ ist .

Für alle $k \geq k_o$ wird dann $n_{k_o-1} > n_{k_o} = n_k$.

Dies besagt : Unter den n ersten Gliedern bleiben n_{k_o} Stellen unbesetzt; nach dem k_o-ten Schritt wird mit den n ersten Stellen keine Aenderung mehr vorgenommen. Da k_o nur von n (und den Eigenschaften der Folge w_1 , w_2 , w_3 , ...) abhängt, so ist die endgültige Besetzung jeder einzelnen Stelle der Folge E' nach endlich vielen Schritten entschieden.

7. Nun zu den Wahrscheinlichkeitseigenschaften der Folge E' !

a) Bei festem k gilt

$$\lim_{n=\infty} \frac{n_k}{n} = w_k + w_{k+1} + w_{k+2} + \cdots \quad , \text{ also insbesondere}$$

$$\lim_{n=\infty} n_k = +\infty$$

In der Tat : Es ist $n_{k+1} = \lambda_k n_k + \delta_k(n)$, wobei

$$\lambda_k = \frac{w_{k+1}+w_{k+2}+w_{k+3}+\cdots}{w_k + w_{k+1}+w_{k+2}+\cdots} \quad , \quad 0 \leq \delta_k(n) < 1 \quad .$$

Daraus folgt für $k \geq 2$ wegen $0 < \lambda_k < 1$:

$$n_k = \lambda_1\lambda_2\cdots\lambda_{k-1}\cdot n + \delta'_{k-1}(n) \quad \text{mit}$$

$$0 \leq \delta'_{k-1}(n) \leq \sum_{\kappa=1}^{k-1} \delta_\kappa(n) < k-1 \quad \text{und}$$

$$\lambda_1\lambda_2\cdots\lambda_{k-1} = w_k + w_{k+1} + w_{k+2} + \cdots \quad .$$

Nun kann man schliessen, dass

$$\lim_{n=\infty} \frac{n_k}{n} = w_k + w_{k+1} + w_{k+2} + \cdots \quad \text{ist.}$$

Diese Behauptung wurde für $k \geq 2$ bewiesen; für $k = 1$ ist sie trivial.

b) Da nach dem kten Schritt die Besetzung von E' insoweit feststeht, als es sich um die Merkmale m_1 , m_2 , $\cdots$, m_k handelt, so ist

$$H_n(E', m_k) = H_{n_k}(\mathfrak{P}^{(k)}, m_k) \quad , \text{ also}$$

$$h_n(E', m_k) = h_{n_k}(F^{(k)}, m_k)\,\frac{n_k}{n} \quad .$$

Geht man hierin bei festem k zur Grenze über, so strebt die rechte Seite gegen

$$\frac{w_k}{w_k+w_{k+1}+w_{k+2}+\cdots}\,(w_k+w_{k+1}+w_{k+2}+\cdots) = w_k \quad .$$

Da die obige Gleichung für alle n gilt, so ist auch

$$\lim_{n=\infty} h_n(E', m_k) = w_k \quad .$$

c) Aus der Definition von $k_o = k_o(n)$ folgt, dass für alle n und alle k $n_k \geq n_{k_o}$ ist. Es gilt also

$$0 \leq \overline{\lim_{n=\infty}} \frac{n_{k_o}}{n} \leq \overline{\lim_{n=\infty}} \frac{n_k}{n} = w_k + w_{k+1} + w_{k+2} + \cdots \quad .$$

Da diese Beziehung für alle k gilt und $w_k+w_{k+1}+w_{k+2}+\cdots$ bei wachsendem k gegen 0 strebt, so existiert der Grenzwert von $\frac{n_{k_o}}{n}$:

$$\lim_{n=\infty} \frac{n_{k_o}}{n} = 0$$

Es sei noch bemerkt, dass diese Behauptung auch mit Hilfe des Satzes 7 bewiesen werden kann.

8. Damit ist gezeigt, dass es eine Folge E' der in Ziffer 2 beschriebenen Art und damit nach Ziffer 3 eine Folge E gibt, welche bezüglich der Merkmale m_1, m_2, m_3, ... die echte Verteilung w_1, w_2, w_3, ... hat. Es bleibt

nur noch der Fall $w_1 + w_2 + w_3 + \ldots < 1$ zu behandeln. Diesen Fall führen wir auf den eben behandelten Fall der echten Verteilung dadurch zurück, dass wir für die Merkmale m' , m_1 , m_2 , m_3 , ... und die echte Verteilung w' , w_1 , w_2 , w_3 , ... mit $w' = 1 - (w_1+w_2+w_3+\ldots)$ eine Folge konstruieren, die bezüglich m' , m_1 , m_2 , m_3 , ... die Verteilung w' , w_1 , w_2 , w_3 , ... hat . Nun ersetzt man das Merkmal m' und zwar durch m_μ, wenn m' an der Stelle μ auftritt. Dadurch ändert sich die relative Häufigkeit des Vorkommens von m_μ höchstens um $\frac{1}{n}$. Der Grenzwert dieser relativen Häufigkeit ändert sich aber nicht, bleibt also w_μ .

Nunmehr können wir den Existenzsatz für W-Folgen mit vorgeschriebener Verteilung aussprechen :

Satz 8. Sind endlich oder abzählbar unendlich viele einander ausschliessende Merkmale m_1 , m_2 , m_3 , ... und ebensoviele Zahlen w_1 , w_2 , w_3 , ... mit $w_\nu \geq 0$ und $w_1 + w_2 + w_3 + \ldots \leq 1$ (diese Summe muss im Fall endlich vieler Merkmale gleich 1 sein) gegeben, so lässt sich eine nur aus den Merkmalen m_1 , m_2 , m_3 , ... aufgebaute E-Folge angeben, die in Bezug auf diese Merkmale die Verteilung w_1 , w_2 , w_3 , ... hat.

§ 12. Z - Folgen.

Wir gehen aus von einer W-Folge $E : E_1 , E_2 , E_3 , \cdots$. Jedem Element der Folge soll eines und nur eines der endlich oder abzählbar unendlich vielen Merkmale $m_1 , m_2 , m_3 , \cdots$ zukommen; die zugehörigen Wahrscheinlichkeiten seien $w_1 , w_2 , w_3 , \cdots$. Unter dem zugehörigen abbrechenden E-Feld $F^{(k)}$ wird das folgende Feld verstanden :

$E_1 , E_2 , E_3 , \cdots$

$E_2 , E_3 , E_4 , \cdots$

$\cdots\cdots\cdots\cdots\cdots\cdots$

$E_k , E_{k+1} , E_{k+2} , \cdots$

Die W-Folge E heisst Z_k-Folge, wenn das E-Feld ein W-Feld in folgendem Sinn ist :

Die Wahrscheinlichkeit dafür, dass in einer Spalte von $F^{(k)}$ in der x_1-ten Zeile m_{r_1} , in der x_2-ten Zeile m_{r_2} , $\cdots$, in der x_ν-ten Zeile m_{r_ν} vorkommen, existiert und ist gleich $w_{r_1} w_{r_2} \cdots w_{r_\nu}$, unabhängig von $1 \leqq x_1 < x_2 < \cdots < x_\nu \leqq k$.

Aus dieser Definition ergibt sich ohne weiteres, dass jede Z_k-Folge eine Z_h-Folge ist, sofern $h \leqq k$ ist.

Mit Hilfe der Z_k-Folge lässt sich nun die Z-Folge definieren:

.VI. Eine E-Folge E heisst Z-Folge, wenn E für jedes $k \geqq 1$ eine Z_k-Folge ist.

Wir werden später zeigen, dass diese Definition sinnvoll ist, d.h., dass zu einer beliebigen Verteilung Z-Folgen angegeben werden können. Vorher soll aber die Definition der Z-Folge bezw. der Z_k-Folge noch näher untersucht werden.

Wir betrachten den Fall der echten Verteilung :
$w_1 + w_2 + w_3 + \ldots = 1$. In diesem Fall können wir nach einem Satz über W-Felder[10] die Definition der Z_k-Folge so aussprechen :
Die E-Folge E heisst Z_k-Folge, wenn das E-Feld $F^{(k)}$ ein W-Feld in folgendem Sinn ist :
Ist $m^{(\nu)}$ ein beliebiges der Merkmale m_1 , m_2 , m_3 , ... , so existiert die Wahrscheinlichkeit dafür, dass in einer Spalte von $F^{(k)}$ in der ersten Zeile m' , in der zweiten Zeile m" , ... , in der ν-ten Zeile $m^{(\nu)}$ vorkommen, und ist gleich $w'w"\ldots w^{(\nu)}$. Dies gilt für alle $1 \leq \nu \leq k$. Es ist noch zu erwähnen, dass, falls alle $w_\nu > 0$ sind, die in den einzelnen Zeilen des Feldes $F^{(k)}$ stehenden Folgen nach der ersten Definition der Z_k-Folge in $F^{(k)}$ unabhängig in Bezug auf jedes Merkmal sind.[11]

Zur weiteren Klärung der Eigenschaften der Z-Folgen benützen wir den Begriff der Nachwirkungsfreiheit.[12]

---

10) vgl. E.Kamke, a.a.O.,S.99/102.

11) vgl. E.Kamke, a.a.O.,S.39/40.

12) vgl. H.Reichenbach, Wahrscheinlichkeitslehre, Leiden 1935, S.134 f.

Wir gehen von einer Folge E : E_1 , E_2 , E_3 , ... aus, welche in Bezug auf die endlich oder abzählbar unendlich vielen einander ausschliessenden Merkmale m_1 , m_2 , m_3 , ... die Verteilung w_1 , w_2 , w_3 , ... mit $w_\nu > 0$ hat. Ferner sei eine endliche Folge von natürlichen Zahlen ρ_1 , ρ_2 , ... , ρ_k gegeben, für die $0 < \rho_1 < \rho_2 < \ldots < \rho_k$ gilt . Betrachtet wird eine Merkmalkombination $m_{r_1} m_{r_2} \ldots m_{r_k}$. Für $\nu = 1, 2, 3, \ldots$ wird untersucht, ob die folgenden Bedingungen gleichzeitig erfüllt sind :

$$E_{\nu-\rho_1} = m_{r_1} \quad ; \quad E_{\nu-\rho_2} = m_{r_2} \quad ; \quad \ldots \quad , \quad E_{\nu-\rho_k} = m_{r_k}$$

Die Zahlen mit dieser Eigenschaft werden, sofern überhaupt solche Zahlen existieren, der Reihe nach mit ν_1 , ν_2 , ν_3 , ... und die Folge E_{ν_1} , E_{ν_2} , E_{ν_3} , ... mit E_N bezeichnet.

E heisst <u>nachwirkungsfrei</u>, wenn für jedes k, jede Folge $0 < \rho_1 < \rho_2 < \ldots < \rho_k$ und jede Merkmalkombination $m_{r_1} m_{r_2} \ldots m_{r_k}$ die zugehörige Teilfolge E_N unendlich und eine W-Folge mit der gleichen Verteilung wie E ist. E_N heisst eine durch Vorgängerauswahl aus E entstandene Folge oder kurz eine Auswahlfolge von E .

Wir wollen nun zeigen, dass für den Fall echter Verteilung und positiver Wahrscheinlichkeiten Z-Folgen und nachwirkungsfreie Folgen identisch sind. Dazu beweisen wir die Sätze 9 und 10 .

Satz 9. Jede Z-Folge mit positiven Wahrscheinlichkeiten ist nachwirkungsfrei.

In der Tat : Ist E eine Z-Folge und $j = \rho_k + 1$, so betrachtet man das zu E gehörende Feld $F^{(j)}$:

E_1 , E_2 , E_3 , ...

E_2 , E_3 , E_4 , ...

.....

E_j , E_{j+1} , E_{j+2} , ...

Aus der Folge $E^{(j)}$ sondert man eine Folge G mit Rücksicht auf das Vorkommen der Merkmalkombination $m_{r_k} m_{r_{k-1}} \cdots m_{r_1}$ in der Verbindungsfolge $E^{(j-\rho_k)} E^{(j-\rho_{k-1})} \cdots E^{(j-\rho_1)}$ aus . G ist gleichzeitig die Auswahlfolge aus E, deren Elemente E die Bedingungen

$E_{\nu-\rho_1} = m_{r_1}$; $E_{\nu-\rho_2} = m_{r_2}$; ... ; $E_{\nu-\rho_k} = m_{r_k}$

in E gleichzeitig erfüllen. Da E eine Z-Folge ist, so ist die Wahrscheinlichkeit $w(G, m_\nu)$ für alle ν unabhängig von dem Auftreten der Merkmalverbindung $m_{r_k} m_{r_{k-1}} \cdots m_{r_1}$ in $E^{(j-\rho_k)} E^{(j-\rho_{k-1})} \cdots E^{(j-\rho_1)}$. Nach der Definition der Unabhängigkeit ist G eine (unendliche) W-Folge mit der gleichen Verteilung wie E.

Die Voraussetzung positiver Wahrscheinlichkeiten ist notwendig. Die Folge E : m_1 , m_2 , m_3 , ... ist eine Z-Folge, bei der alle Wahrscheinlichkeiten 0 sind. E ist aber nicht nachwirkungsfrei.

Satz 10. Jede nachwirkungsfreie Folge mit echter Verteilung ist eine Z-Folge.

In der Definition der Nachwirkungsfreiheit setzt man zum Beweis dieses Satzes $\rho_1 = 1$, $\rho_2 = 2$, ... , $\rho_k = k$. Für $k = 1$ besagt dann die Definition der Nachwirkungsfreiheit, dass die beiden ersten Zeilen E' und E" des Feldes $F^{(k)}$ in Bezug auf jedes der Merkmale m_1 , m_2 , m_3 , ... unabhängig sind. Nach dem Multiplikationssatz ist also

1) $w(E'E",m'm") = w(E',m') w(E",m")$,

wobei m' und m" jedes der Merkmale m_1 , m_2 , m_3 , ... bedeuten können.

Für beliebiges k sei bereits bewiesen :

2) $w(E'E"...E^{(k)},m'm"...m^{(k)}) = w(E',m')w(E",m")...w(E^{(k)},m^{(k)})$

Durch Vorgängerauswahl mittels $m'm"...m^{(k)}$ entsteht aus E eine Teilfolge G, die in Bezug auf die Merkmale m_1 , m_2 , m_3 , ... die Verteilung w_1 , w_2 , w_3 , ... hat. Nach der Definition der Unabhängigkeit ist dann $w(E^{(k+1)},m^{(k+1)})$ unabhängig von dem Auftreten der Merkmalkombination $m'm"...m^{(k)}$ in $E'E"...E^{(k)}$. Nach dem Multiplikationsgesetz für unabhängige Folgen ergibt sich bei Berücksichtigung von (2) , wenn $w^{(i)} = w(E^{(i)},m^{(i)})$ gesetzt wird :

3) $w(E'E"...E^{(k+1)},m'm"...m^{(k+1)}) = w'w"...w^{(k+1)}$.

Aus (1) und (3) folgt, dass (2) für jedes k gilt. Die Folge E ist also, da $w_1+w_2+w_3+... = 1$ vorausgesetzt wird, eine Z-Folge.

Die Sätze 9 und 10 erlauben noch eine Folgerung für Folgen mit echter Verteilung. Nennt man eine Folge, für die die Voraussetzungen der Nachwirkungsfreiheit wenigstens für $\rho_i = i$ zutreffen, " nachwirkungsfrei ohne Lücken " und eine im Sinne der Definition nachwirkungsfreie Folge " nachwirkungsfrei mit Lücken " , so ergibt der Beweis des Satzes 10, dass jede Folge mit echter Verteilung, die nachwirkungsfrei ohne Lücken ist, eine Z-Folge und damit nach Satz 9 eine Folge ist, die nachwirkungsfrei mit Lücken ist. Da die Umkehrung trivial ist, so genügt es bei echter Verteilung die Definition der Nachwirkungsfreiheit für $\rho_i = i$ auszusprechen.

§ 13. Sätze über Z - Folgen.

Wir betrachten eine Folge E mit den Elementen m und $\bar{m}$. Es sei E eine X-Folge, $[a,b]$ das kleinste Intervall, in dem die W-Bänder von E und von allen $E^{(k,r)}$ enthalten sind. Wir zeigen nun, dass Z-Folgen und X-Folgen mit $a = b$ identisch sind. Dies ergibt sich ohne weiteres daraus, dass die Auswahlfolgen E_N mit den Folgen $E^{(k,r)}$ übereinstimmen. Z-Folgen sind also X-Folgen, bei denen die Folge selbst und alle Auswahlfolgen W-Folgen sind. Die gleiche Ueberlegung führt zu dem Ergebnis, dass Z-Folgen ein Spezialfall der Poisson'schen X-Folgen sind.

Aus diesem Gedankengang ergibt sich nach § 10 ohne weiteres der Unempfindlichkeitssatz für Z-Folgen :

atz 11. Ist $E : E_1 , E_2 , E_3 , \dots$ eine Z-Folge und nimmt man in E Streichungen oder Abänderungen der mit der Grenzhäufigkeit 0 auftretenden Elemente $E_{\nu_1} , E_{\nu_2} , E_{\nu_3} , \dots$ vor, so ist die dadurch entstehende Folge wieder eine Z-Folge mit der gleichen Verteilung wie E.

Weiter folgt aus dem Zusammenhang von Poisson'schen Verteilungssystemen und Z-Folgen die Gültigkeit der beiden Aussagen des Bernoullischen Gesetzes auch für Z-Folgen. Darüber hinaus lässt sich für Z-Folgen die Newton'sche Formel beweisen. Dazu benötigen wir einen Hilfssatz.

fssatz 5. Gegeben sei ein E-Feld

$E' : E'_1 , E'_2 , E'_3 , \dots$

$E'' : E''_1 , E''_2 , E''_3 , \dots$

$E''' : E'''_1 , E'''_2 , E'''_3 , \dots$

.....

In Bezug auf das Merkmal m gelte für jedes beliebige Indexsystem $1 \leq r_1 < r_2 < r_3 < \dots < r_\nu$

1) $w(E^{(r_1)} E^{(r_2)} \dots E^{(r_\nu)}, m \dots m) = w^\nu$.

Haben alle Elemente, denen das Merkmal m nicht zukommt, das Merkmal $\bar{m}$, so gilt die Beziehung (1) in analoger Weise auch für jede beliebige Kombination von m und $\bar{m}$, nämlich:

Ist $w_{m_\lambda} = w$ für $m_\lambda = m$, und

$w_{m_\lambda} = 1-w$ für $m_\lambda = \bar{m}$, so gilt

$$w(E^{(r_1)}E^{(r_2)}\dots E^{(r_\nu)}, m_1 m_2 \dots m_\nu) = w_{m_1} w_{m_2} \dots w_{m_\nu} .$$

Zum Beweis möge λ einer der Indices 1, 2, ... , ν sein. Auf Grund des Divisionsgesetzes sagt (1) aus : Die aus $E^{(r_\lambda)}$ mit Rücksicht auf das Vorkommen von mm...m in $E^{(r_1)}E^{(r_2)}\dots E^{(r_\lambda-1)}E^{(r_\lambda+1)}\dots E^{(r_\nu)}$ ausgesonderte Teilfolge ist in Bezug auf m eine W-Folge mit der Wahrscheinlichkeit w und also in Bezug auf $\bar{m}$ eine W-Folge mit der Wahrscheinlichkeit 1-w . Nach Satz 3 bei Kamke Seite 38 existiert also die folgende Wahrscheinlichkeit und hat den Wert

$$w(E^{(r_1)}\dots E^{(r_\lambda)}\dots E^{(r_\nu)}, m\dots m\bar{m}\bar{m}\dots m) = w^{\nu-1}(1-w) .$$

Durch wiederholte Anwendung dieser Betrachtung kommt man zu der Behauptung des Hilfssatzes 5.

Wir kehren nun zu der aus den Merkmalen m und $\bar{m}$ aufgebauten Z-Folge E zurück. Die Wahrscheinlichkeit dafür, dass unter k aufeinanderfolgenden Elementen der Z-Folge E genau r Elemente m vorkommen, wird mit $W_{k,r}$ bezeichnet. Das Merkmal " genau r mal m " ist die Mischung aus den $\binom{k}{r}$ verschiedenen Merkmalverbindungen , die aus r Merkmalen m und k-r Merkmalen $\bar{m}$ bestehen. Es ist also zu untersuchen, welche Wahrscheinlichkeit für das Auftreten einer solchen Merkmalverbindung gilt. Dazu

betrachtet man die Verbindungsfolge $E'E''\ldots E^{(k)}$ und wendet den eben bewiesenen Hilfssatz an. Für eine bestimmte Merkmalverbindung ergibt sich die Wahrscheinlichkeit $w^r\ (1-w)^{k-r}$; nach dem Additionssatz folgt daraus die Newton'sche Formel :

$$W_{k,r} = \binom{k}{r} w^r\ (1-w)^{k-r}$$

Für Z-Folgen gelten daher die Bernoullischen Gesetze in der üblichen Form.

§ 14. Satz über die Auswahlfolge einer Z - Folge.

Nach der Definition der Nachwirkungsfreiheit ist die Auswahlfolge einer Z-Folge eine W-Folge. Wir wollen nun zeigen, dass in der Definition der Z-Folge wesentlich mehr steckt, nämlich :

atz 12. Sind die Wahrscheinlichkeiten w_ν positiv und $\sum w_\nu = 1$, so ist jede durch Vorgängerauswahl aus einer Z-Folge E gebildete Folge F wieder eine Z-Folge.

Es genügt zu zeigen, dass die Folge G, die aus F durch Vorgängerauswahl mittels der beliebigen Kombination $m_1 m_2 \ldots m_k$ entsteht, die gleiche Verteilung wie E hat. Die Kombination mittels deren F aus E hervorgeht, sei C. Die Kombinationen $m_1 m_2 \ldots m_k$ und C halten wir im folgenden fest.

Es sei K eine Kombination mit folgenden Eigenschaften :

1. C kommt in K genau (k+1)-mal vor.
2. C steht am Anfang und am Ende von K.
3. Nach dem ersten C in K steht m_1, nach dem zweiten C m_2, ... , nach dem k-ten C m_k .

Jeder Kombination K wird die aus E mittels dieser Kombination ausgewählte Teilfolge A zugeordnet : A = E(K) . Ueber diese Folgen A lässt sich zeigen :

1. Sind A' und A" nach zwei verschiedenen Kombinationen K' und K" ausgewählt, so sind sie elementefremd.

   Würden nämlich A' und A" gemeinsame Elemente, z.B. E_σ enthalten, so müsste E_σ sowohl die Vorgängerkombination K' als auch die Vorgängerkombination K" haben. Dann müsste etwa K' in K" enthalten sein, was zur Folge hätte, dass K" mehr als k+1 Kombinationen C enthielte. Das widerspricht aber den Voraussetzungen über die K.

2. Jedes Element von G kommt in einer Folge A vor.

   m' sei ein Element von G; m' hat also in F die Vorgänger $m_1 m_2 \ldots m_k$. In E müssen also die Kombinationen Cm_1 , Cm_2 , ... , Cm_k und Cm' vorkommen. Damit die Elemente m_1 , m_2 , ... , m_k und m' in F unmittelbar aufeinanderfolgen, darf in und zwischen Cm_1 und Cm_2 kein drittes C liegen, es sei

denn, dass es in Cm_2 enthalten ist ; dasselbe muss für Cm_2 und Cm_3 , ... , Cm_k und Cm' gelten. Dies ist aber in der Definition von K verlangt. Tatsächlich hat also jedes Element von G eine Vorgängerkombination von der Form K, gehört demnach zu einem **A**.

3. Jedes Element eines **A** gehört zu G.

m' sei ein Element irgend eines **A** und hat also in E eine Vorgängerkombination von der Form K. Wählt man aus **E** die Elemente aus, die die Vorgänger C haben, so gelangt man zum Abschnitt $m_1m_2...m_km'$ von F. m' hat in F die Vorgängerkombination $m_1m_2...m_k$ und gehört somit zu G.

4. Die Kombinationen K lassen sich ordnen : K_1 , K_2 , K_3 , ...

Die Kombinationen werden so in Gruppen eingeteilt, dass in die Gruppe mit der Nummer n diejenigen Kombinationen kommen, die aus genau n Gliedern bestehen. Innerhalb jeder Gruppe lassen sich die Kombinationen in Klassen einteilen : Die erste Klasse enthält alle Kombinationen, die nur aus m_1 bestehen.
Die zweite Klasse enthält alle Kombinationen, die nur aus m_1 und m_2 bestehen, aber nicht zur ersten Klasse gehören.
.....
Die n-te Klasse enthält alle Kombinationen, die nur aus m_1 , m_2 , ... , m_n bestehen, aber nicht zu einer der vorhergehenden Klassen gehören.

.....

Die Anzahl der zu einer bestimmten Klasse gehörenden Kombinationen ist endlich. Die Kombinationen einer und derselben Gruppe lassen sich also durchzählen, wenn man zunächst die Kombinationen der ersten Klasse, dann diejenigen der zweiten Klasse usf. nimmt. Wir können also die Kombinationen in der Weise nummerieren, dass ein Index die Gruppe und ein zweiter Index die Nummer innerhalb der Gruppe angibt. Da stets nur endlich viele Kombinationen vorhanden sind, deren Indices eine bestimmte Summe haben, so wird eine Ordnung der Kombinationen erreicht, wenn man zuerst die Kombinationen mit der Indices-Summe 2, dann diejenigen mit der Indices-Summe 3 usf. nimmt. Dass es vorkommen wird, dass in einer Klasse oder Gruppe keine Kombination vorhanden ist, stört die Abzählung nicht.

5. Wegen $A = E(K)$ bewirkt die eben durchgeführte Abzählung der K eine solche der A. Man kann deshalb schreiben : $A_1 , A_2 , A_3 , \dots$.

6. Die Wahrscheinlichkeit dafür, dass ein Element von E zu A_r gehört, ist bekannt; es ist dies die wegen der Voraussetzung, dass E Z-Folge ist, sicher existierende Wahrscheinlichkeit für das Auftreten der Merkmalkombination K_r in E . Sie möge mit p_r bezeichnet werden; es ist $p_r > 0$. Da die Folgen A_r elementefremd sind, so sind für jedes r die Teilsummen

$$s_r = p_1 + p_2 + \dots + p_r \leqq 1 ,$$

also beschränkt; ferner ist $s_{r+1} > s_r$. Es existiert also $\sum_{r=1}^{\infty} p_r$ und es ist $0 < \sum_{r=1}^{\infty} p_r \leq 1$.

7. Mit p bezeichnen wir die Wahrscheinlichkeit dafür, dass ein Element von E zu G gehört; es ist, falls p überhaupt existiert, nach dem Additionsgesetz $p \geq \sum_{r=1}^{\infty} p_r$. Wir werden später zeigen, dass p existiert und $p = \sum_{r=1}^{\infty} p_r$ gilt. Dies soll aber für das folgende schon benützt werden.

8. Unter den n ersten Elementen von E treten n_r Elemente auf, die zu A_r gehören, und n' Elemente, die zu G gehören. Die relative Häufigkeit für das Auftreten von Elementen von A_r in G ist dann $\frac{n_r}{n'}$. Da $\lim \frac{n'}{n} = p > 0$ ist, so wird

$$\lim_{n'=\infty} \frac{n_r}{n'} = \frac{\lim_{n=\infty} \frac{n_r}{n}}{\lim_{n=\infty} \frac{n'}{n}} = \frac{p_r}{p} \quad .$$

$q_r = \frac{p_r}{p}$ ist also die Wahrscheinlichkeit dafür, dass ein Element von G auch ein Element von A_r ist. Wegen

$$\sum_{r=1}^{\infty} p_r = p \quad \text{ist} \quad \sum_{r=1}^{\infty} q_r = 1 \quad .$$

9. Die Folgen A_r sind Auswahlfolgen der Z-Folge E und haben als solche die gleiche Verteilung wie E. Wir zeigen, dass auch G die gleiche Verteilung wie E hat.

Zum Beweis dieser Behauptung betrachtet man das Merkmal m mit der zugehörigen Wahrscheinlichkeit w und bezeichnet mit n_r die Anzahl derjenigen unter den n ersten Gliedern von G, welche Glieder von A_r sind. Es ist $n = n_1 + n_2 + n_3 + \dots$. Für die entsprechenden Anzahlen und Häufigkeiten für m ergibt sich :

$$H_n(G) = H_{n_1}(A_1) + H_{n_2}(A_2) + \dots + H_{n_\nu}(A_\nu) + H_{n_{\nu+1}}(A_{\nu+1}) + \dots$$

$$h_n(G) = h_{n_1}(A_1)\,\frac{n_1}{n} + h_{n_2}(A_2)\,\frac{n_2}{n} + \dots + h_{n_\nu}(A_\nu)\,\frac{n_\nu}{n} + h_{n_{\nu+1}}(A_{\nu+1})\,\frac{n_{\nu+1}}{n} + \dots$$

$$h_n(G) - w = \left(h_{n_1}(A_1) - w\right)\frac{n_1}{n} + \left(h_{n_2}(A_2) - w\right)\frac{n_2}{n} + \dots + \left(h_{n_\nu}(A_\nu) - w\right)\frac{n_\nu}{n} + \left(h_{n_{\nu+1}}(A_{\nu+1}) - w\right)\frac{n_{\nu+1}}{n} + \dots$$

Es folgt dann, da $\left|h_{n_s}(A_s) - w\right| \leq 1$ ist :

$$\left|h_n(G) - w\right| \leq \left|h_{n_1}(A_1) - w\right|\frac{n_1}{n} + \dots + \left|h_{n_\nu}(A_\nu) - w\right|\frac{n_\nu}{n} + \left(1 - \frac{n_1+n_2+\dots+n_\nu}{n}\right)$$

Nun wählt man r so gross, dass $q_1+q_2+\ldots+q_r > 1-\epsilon$ wird. Ein solches r wird festgehalten und dann N so gewählt, dass für alle $n \geq N$ gleichzeitig folgende Bedingungen erfüllt sind :

$$\left| \frac{n_1}{n} - q_1 \right| < \frac{\epsilon}{r} \quad ; \qquad \left| h_{n_1}(A_1) - w \right| < \epsilon$$

.....

$$\left| \frac{n_r}{n} - q_r \right| < \frac{\epsilon}{r} \quad ; \qquad \left| h_{n_r}(A_r) - w \right| < \epsilon$$

Jetzt wird

$$\left| h_n(G) - w \right| < \epsilon \; \frac{n_1+n_2+\ldots+n_r}{n} + \left(1 - \frac{n_1+n_2+\ldots+n_r}{n} \right)$$

$$< \epsilon + \left(1 - \left(q_1+q_2+\ldots+q_r \right) + \epsilon \right) < 3\epsilon .$$

Die Folge G hat also die Verteilung w_1 , w_2 , w_3 , ...

10. Wir haben noch den Beweis für $p = \sum_{r=1}^{\infty} p_r$ nachzutragen.

Zu diesem Beweis benötigen wir einige Eigenschaften der Z-Folgen, die zunächst abgeleitet werden mögen.

Gegeben sei eine Z-Folge E, die in Bezug auf die einander ausschliessenden Merkmale m_1 , m_2 , m_3 , ... die echte Verteilung w_1 , w_2 , w_3 , ... hat. Man betrachtet nun die aus den Merkmalen m_1 , m_2 , m_3 , ... aufgebauten Merkmalkombinationen $C^{(s)}$ von fester Länge s. (Unter der Länge der Merkmalkombination verstehen wir die Anzahl der in ihr

enthaltenen gleichen oder verschiedenen Merkmale.)
Diese Kombinationen $C^{(s)}$ lassen sich ordnen; man kann also $C_\nu^{(s)}$ schreiben. Dann lautet die Behauptung :

$$\sum_{\nu=1}^{\infty} w(F^{(s)}, C_\nu^{(s)}) = 1$$

Unter $F^{(s)}$ möge hier die Verbindungsfolge aus den ersten s Zeilen des zu E gehörenden Feldes verstanden werden.
Für $s = 1$ ist die Behauptung in der Voraussetzung

$\sum_{\nu=1}^{\infty} w_\nu = 1$ enthalten; sie sei bereits für $s \leq \sigma$ bewiesen. Wir betrachten nun die Kombinationen $C_\mu^{(\sigma+1)}$.
Diese Kombinationen ordnen wir um; anstelle von $C_\mu^{(\sigma+1)}$ schreiben wir $C_{\lambda,\nu}^{*(\sigma+1)}$, wenn $C_\mu^{(\sigma+1)}$ zunächst die Kombination $C_\lambda^{(\sigma)}$ und dann das Merkmal m_ν enthält; λ und ν laufen über alle natürlichen Zahlen. Aus der Definition der Z-Folgen ergibt sich

$$w(F^{(\sigma+1)}, C_{\lambda,\nu}^{*(\sigma+1)}) = w(F^{(\sigma)}, C_\lambda^{(\sigma)}) w_\nu$$

Nun ist

$$\sum_{\mu=1}^{\infty} w(F^{(\sigma+1)}, C_\mu^{(\sigma+1)}) = \sum_{\lambda,\nu} w(F^{(\sigma+1)}, C_{\lambda,\nu}^{*(\sigma+1)})$$

$$= \sum_{\lambda,\nu} w(F^{(\sigma)}, C_\lambda^{(\sigma)}) w_\nu$$

$$= \sum_{\lambda=1}^{\infty} \left\{ w(F^{(\sigma)}, C_\lambda^{(\sigma)}) \sum_{\nu=1}^{\infty} w_\nu \right\} = 1$$

11. Ist M eine Mischung von endlich oder abzählbar unendlich vielen Kombinationen $C_\mu^{(s)}$ (s fest) , so existiert nach § 11 Satz 7 die folgende Wahrscheinlichkeit und hat den Wert

$$w(F^{(s)}, M) = \sum w(F^{(s)}, C^{(s)}) \quad .$$

12. Nunmehr können wir zeigen, dass p existiert und

$p = \sum_{r=1}^{\infty} p_r$ ist. Nach Satz 5 bei Kamke Seite 26 gilt

$\underline{p} \geq \sum_{r=1}^{\infty} p_r$. Dabei ist $\underline{p}$ die untere Grenzhäufigkeit dafür, dass ein Element von E zu G gehört. Bedeutet $\overline{p}$ die entsprechende obere Grenzhäufigkeit, so genügt es,

$\overline{p} \leq \sum_{r=1}^{\infty} p_r$ zu beweisen.

p(s) sei die Wahrscheinlichkeit dafür, dass ein Element von E ein K_r von der Länge $\leq s$ als Vorgänger hat. Nach Ziffer 11 existiert p(s) und es ist $p(s) = \sum p_{r'}$, wobei nur über diejenigen r' summiert wird, für welche $K_{r'}$ eine Länge $\leq s$ hat. Sicher ist also

$$p(s) \leq \sum_{r=1}^{\infty} p_r \quad .$$

Mit $\overline{q}(s)$ wird die obere Grenzhäufigkeit dafür bezeichnet, dass ein Element von E einen Vorgänger K_r von einer Länge $> s$ hat. Nun ist

$$\overline{p} = p(s) + \overline{q}(s) ,$$

da jedes Element von E einen Vorgänger K_r hat. Wir zeigen nachher, dass $\lim \overline{q}(s) = 0$ ist. Es ist dann

$$\overline{p} \leq \sum_{r=1}^{\infty} p_r + \overline{q}(s) .$$

Da diese Beziehung für jedes s gilt, so gilt auch

$$\overline{p} \leq \sum_{r=1}^{\infty} p_r + \lim_{s=\infty} \overline{q}(s) \quad \text{und damit}$$

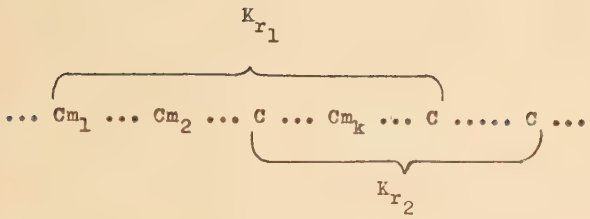
$$\overline{p} \leq \sum_{r=1}^{\infty} p_r .$$

13. Wir zeigen nun noch

$$\overline{q}(s) \leq \frac{k + c - 1}{s}$$

Dabei bedeutet c die Länge von C .

In E können sich zwei K_r nur so stückweise überdecken (etwa K_{r_2} von rechts her ein K_{r_1}) , dass ein C von K_{r_1} das erste C von K_{r_2} ist oder dass das letzte C von K_{r_1} und das erste C von K_{r_2} sich stückweise überdecken :



Mit E_{n_1} bezeichnen wir ein Element von E, das K_{r_1} als Vorgänger hat. Die K_{r_2} können beginnen bei den vor m_2, vor m_3, ... , vor m_k stehenden C, ferner bei jedem der c Elemente des letzten C von K_{r_1} ; es gibt also höchstens $k - 1 + c$ Vorgängerkombinationen K_{r_2} , also höchstens $k - 1 + c$ Elemente nach E_{n_1} , die eine Vorgängerkombination K_r haben, die sich stückweise mit K_{r_1} überdeckt.

Nun wählen wir eine beliebige Vorgängerkombination K_{r_1} und betrachten die ersten s Elemente nach E_{n_1} . Unter diesen s Elementen befinden sich Elemente mit einer Vorgängerkombination K_r von einer Länge $>$ s nur dann, wenn dieses K_r und K_{r_1} sich stückweise überdecken. Dies kann aber auf höchstens $k + c - 1$ Elemente zutreffen. Sicher ist also

$$\bar{q}(s) \leq \frac{k + c - 1}{s} .$$

Da k und c feste Grössen sind, so folgt

$$\lim_{s=\infty} \overline{q}(s) = 0 \ .$$

Damit ist der Beweis beendet.

§ 15. Der Existenzsatz für Z - Folgen.

1. Es sei $p \geq 2$, $k \geq 2$. Aus den p verschiedenen Merkmalen $m_1, m_2, \ldots, m_p$ lassen sich p^k Kombinationen zur k-ten Klasse bilden. Diese p^k Kombinationen können nach folgenden Vorschriften geordnet werden :

(α) Die erste Kombination ist $m_1 m_1 \ldots m_1$.

(β) Auf die Kombination $M_1 M_2 \ldots M_k$ folgt die Kombination $M_2 M_3 \ldots M_k M$, wobei $M = m_j$ und j der grösste Index derart ist, dass die Kombination $M_2 M_3 \ldots M_k m_j$ bisher noch nicht vorgekommen ist.

Es wird gezeigt, dass die Anordnung von zyklischer Bauart ist und dass alle Kombinationen vorkommen.

a) Spätestens nach p^k Schritten muss das Verfahren abbrechen. Die letzte Kombination sei $K^* = M_1 M_2 \ldots M_k$. Das bedeutet nach (β), dass die Kombinationen $K_i = M_2 M_3 \ldots M_k m_i$ ($i = 1, 2, \ldots, p$) schon angeschrieben sind. Die Vorgänger der K_i, sofern solche vorhanden sind, können sein :

$K_j' = m_j M_2 M_3 \cdots M_k$ ($j = 1, 2, \cdots, p$) . Da sie aber vor K^* liegen und also von K^* verschieden sind, ist $m_j \neq M_1$. Es stehen nur p-1 Kombinationen K_j' zur Verfügung. Da die K_i alle voneinander verschieden sind, so gilt dies auch für die Vorgänger je zweier der K_i . Eines der K_i kann demnach keinen Vorgänger haben, muss also die erste Kombination der Anordnung sein. Wegen (α) ist dies nur für K_1 möglich; es ist

$M_2 = M_3 = \cdots = M_k = m_1$.

Daraus folgt der zyklische Charakter.

b) Sicher kommen folgende Kombinationen vor :

$K_i = m_1 m_1 \cdots m_1 m_i$ ($i = 1, 2, \cdots, p$) wegen a)

$m_j m_1 \cdots m_1$ ($j = 1, 2, \cdots, p$) als Vorgänger von K_i ($i = 2, 3, \cdots, p$) bzw. als letztes Element.

Damit müssen aber bei festem $j \neq 1$ auch vorhanden sein :

$m_j m_1 \cdots m_1 m_q$ ($q = 1, 2, \cdots, p$) nach (β) , also auch deren p Vorgänger

$m_{j'} m_j m_1 \cdots m_1$ ($j' = 1, 2, \cdots, p$) .

Der Schluss lässt sich wiederholen. Man kommt auf diese Weise zu $m_{j^{(k-1)}} m_{j^{(k-2)}} \cdots m_{j''} m_{j'} m_j$. Da hierin

$j, j', j'', \cdots, j^{(k-1)}$ unabhängig voneinander alle Werte $1, 2, 3, \cdots, p$ annehmen können, so ist gezeigt, dass tatsächlich alle p^k Kombinationen Verwendung finden.

Der so bewiesene Anordnungssatz besagt nun : Es gibt eine E-Folge A von p^k+k-1 Gliedern, bestehend aus den Merkmalen $m_1, m_2, \dots, m_p$:

$$E_1 E_2 \dots E_{p^k} E_{p^k+1} \dots E_{p^k+k-1}$$

Diese Folge A hat die Eigenschaft, dass jede Kombination zur k-ten Klasse durch genau einen Abschnitt von k aufeinanderfolgenden Gliedern dargestellt wird und dass die ersten k-1 und die letzten k-1 Glieder von A der Reihe nach übereinstimmen. Die ersten p^k Glieder von A heissen eine Grundfolge $G_{k,p}$; p^k ist die Periodenlänge. Im folgenden wollen wir nun diejenige spezielle Grundfolge verwenden, deren k erste Glieder sämtlich gleich m_1 sind.

2. Bildet man mit einer und derselben speziellen Grundfolge $G_{k,p}$ die periodische E-Folge

(B) $G_{k,p}, G_{k,p}, \dots$,

so ist, wie sich aus den Eigenschaften der Grundfolge und aus der Periodizität von (B) ergibt, (B) in Bezug auf $m_1, m_2, \dots, m_p$ eine Z_k-Folge mit Gleichverteilung. Setzt man in $G_{k,p}$

$$m_1 = m_2 = \dots = m_{p_1} = M_1$$

$$m_{p_1+1} = m_{p_1+2} = \dots = m_{p_1+p_2} = M_2$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$m_{p_1+p_2+\dots+p_{q-1}+1} = \dots = m_p = M_q$$

mit $p_1+p_2+ \dots +p_q = p$,

so sind die k ersten Glieder von $G_{k,p}$ gleich M_1 und aus (B) entsteht auf Grund der Additionssätze der Wahrscheinlichkeitsrechnung eine Z_k-Folge mit der rationalen Verteilung

$$w(M_1) = \frac{p_1}{p} \quad , \quad w(M_2) = \frac{p_2}{p} \quad , \dots , \quad w(M_q) = \frac{p_q}{p} \quad .$$

Demnach lässt sich zu beliebig vorgegebener rationaler Verteilung in Bezug auf M_1, M_2, ... , M_q stets eine Z_k-Folge mit dieser Verteilung angeben.

3. Nun sei eine beliebige (irrationale) Verteilung

w_1, w_2, w_3, ... mit $\sum_{q=1}^{\infty} w_q = 1$ in Bezug auf

m_1, m_2, m_3, ... gegeben. Die Zahlen w_q approximieren wir nach folgender Vorschrift durch rationale Zahlenfolgen $w_q^{(k)}$:

$w_1^{(1)} = 1$.

Aus den bereits gewonnenen Zahlen $w_q^{(k)}$ ($q = 1,2,\dots,k$) bilden wir die Zahlen $w_q^{(k+1)}$. Für $q = 1,2,\dots,k$ wählen wir $w_q^{(k+1)}$ rational in $w_q < w_q^{(k+1)} < w_q^{(k)}$ so nahe an w_q, dass

$$\sum_{q=1}^{k} \left(w_q^{(k+1)} - w_q \right) < w_{k+2} + w_{k+3} + \dots$$

wird. Ferner setzen wir dann $w_{k+1}^{(k+1)} = 1 - \sum_{q=1}^{k} w_q^{(k+1)}$.

Dann ist

$$w_{k+1}^{(k+1)} = \sum_{q=1}^{\infty} w_q - \sum_{q=1}^{k} w_q^{(k+1)}$$

$$= \sum_{q=1}^{k} \left(w_q - w_q^{(k+1)} \right) + w_{k+1} + w_{k+2} + w_{k+3} + \cdots$$

oder

$$w_{k+1}^{(k+1)} - w_{k+1} = w_{k+2} + w_{k+3} + \cdots - \sum_{q=1}^{k} \left(w_q^{(k+1)} - w_q \right)$$

$$> 0 .$$

Da die $w_q^{(k)}$ monoton fallende, beschränkte Zahlen sind, so existiert ein Grenzwert. Wegen der Konvergenz der Reihe

$\sum_{q=1}^{\infty} w_q$ und wegen (1) gilt

(2) $$\lim_{k=\infty} w_q^{(k)} = w_q .$$

Nun setzt man

(3) $$\delta_k = \max_{\substack{1 \le r_1 < r_2 < \cdots < r_\nu \\ 1 \le \nu \le k}} \left| w_{r_1}^{(k)} w_{r_2}^{(k)} \cdots w_{r_\nu}^{(k)} - w_{r_1} w_{r_2} \cdots w_{r_\nu} \right|$$

Zunächst zeigen wir, dass für positive Zahlen a_ν' und a_ν , die kleiner als 1 sind, die Ungleichung (4) gilt :

(4) $$\left| a_1' a_2' \cdots a_\nu' - a_1 a_2 \cdots a_\nu \right| < \left| a_1' - a_1 \right| + \left| a_2' - a_2 \right| + \cdots + \left| a_\nu' - a_\nu \right|$$

Für $\nu = 1$ ist die Behauptung trivial. Sie sei für

$\nu = n$ bereits bewiesen. Dann folgt

$$\begin{aligned}|a_1'\cdots a_{n+1}' - a_1\cdots a_{n+1}| &= |a_1'\cdots a_n'a_{n+1}' - a_1\cdots a_n a_{n+1}' \\ &\quad + a_1\cdots a_n a_{n+1}' - a_1\cdots a_n a_{n+1}| \\ &\leq a_{n+1}'\,|a_1'\cdots a_n' - a_1\cdots a_n| \\ &\quad + a_1\cdots a_n\,|a_{n+1}' - a_{n+1}| \\ &< |a_1'-a_1| + \cdots + |a_n'-a_n| + |a_{n+1}'-a_{n+1}|\end{aligned}$$

Wendet man (4) auf δ_k an, so wird

$$\begin{aligned}\delta_k &< (w_1^{(k)}-w_1) + (w_2^{(k)}-w_2) + \cdots + (w_k^{(k)}-w_k) \\ &= w_{k+1} + w_{k+2} + w_{k+3} + \cdots\end{aligned}$$

Wegen der Konvergenz der Reihe $\sum_{q=1}^{\infty} w_q = 1$ strebt δ_k mit wachsendem k gegen 0 . Nun setzt man $\epsilon_k = \mathrm{Max}(\delta_k, \delta_{k+1}, \delta_{k+2}, \cdots)$. Dann ist $\lim_{k=\infty} \epsilon_k = 0$ und zwar strebt ϵ_k monoton gegen 0 .

Nach Ziffer 2 wird nun zu jeder Verteilung $w_1^{(k)}, w_2^{(k)}, \cdots, w_k^{(k)}$ eine Z_k-Folge $F^{(k)}$ für $k = 1, 2, 3, \cdots$ konstruiert. p_k sei die Periodenlänge von $F^{(k)}$.

Jetzt bestimmt man Hilfszahlen $M_1, M_2, M_3, \cdots$ derart, dass

) jede Zahl M_k ein ganzes Vielfaches von p_k ,

(β) $$\frac{M_1\epsilon_1 + M_2\epsilon_2 + \ldots + M_k\epsilon_k}{M_1 + M_2 + \ldots + M_k} < 2\,\epsilon_k$$

(γ) $$\frac{p_{k+1}}{M_1 + M_2 + \ldots + M_k} < \epsilon_k$$

(δ) $2\,(M_1 + M_2 + \ldots + M_{k-1}) < M_1 + M_2 + \ldots + M_k$ wird,

was man offenbar schrittweise erreichen kann.

Zur Abkürzung führen wir ein :

$M_1 + M_2 + \ldots + M_k = \mu_k$

$w_{r_1} w_{r_2} \ldots w_{r_v} = \overline{w}$

$w^{(k)}_{r_1} w^{(k)}_{r_2} \ldots w^{(k)}_{r_v} = \overline{w}^{(k)}$

Die gesuchte E-Folge E wird aus den Folgen $F^{(k)}$ zusammengestückelt. Der Reihe nach werden angeschrieben :

Die M_1 ersten Elemente von F' , dann

die M_2 ersten Elemente von F" , dann

..... , dann

die M_k ersten Elemente von $F^{(k)}$ usf.

Um zu zeigen, dass diese Folge E eine Z-Folge ist, genügt es zu zeigen, dass E für jedes k eine Z_k-Folge ist. Vorgegeben wird eine Folge von Zahlen $1 \leq x_1 < x_2 < \ldots < x_v$ und eine Folge entsprechender Zahlen $1 \leq r_1 < r_2 < \ldots < r_v$.

Es sei $k > \mathrm{Max}(x_\nu, r_\nu)$, aber sonst beliebig gewählt. k werde festgehalten. In dem zu E gehörenden abbrechenden $F^{(k)}$-Feld betrachten wir die x_1-te, x_2-te, ... , x_ν-te Zeile und fassen die Merkmalverbindung $m_{r_1} m_{r_2} \dots m_{r_\nu}$ ins Auge. Durch Streichung der ersten μ_{k-1} Glieder entsteht aus E die Restfolge $E^{[k]}$. Nun ist nach dem Unempfindlichkeitssatz für W-Folgen die Wahrscheinlichkeit für das Auftreten der Merkmalverbindung $m_{r_1} m_{r_2} \dots m_{r_\nu}$ in den betrachteten Zeilen des zu $E^{[k]}$ gehörenden abbrechenden Feldes die gleiche wie in den entsprechenden Zeilen des zu E gehörenden Feldes. Es genügt also zu zeigen, dass es zu jedem $\epsilon > 0$ ein N derart gibt, dass für alle $n \geq N$

5) $$\left| h_n\left(E^{[k]}, \frac{x_1 x_2 \dots x_\nu}{r_1 r_2 \dots r_\nu} \right) - \bar{w} \right| < \epsilon$$

ist.

Für den Beweis benötigen wir die Ungleichung

6) $$\frac{\epsilon_k M_k + \epsilon_{k+1} M_{k+1} + \dots + \epsilon_\lambda M_\lambda}{M_k + M_{k+1} + \dots + M_\lambda} \leq \frac{\epsilon_1 M_1 + \epsilon_2 M_2 + \dots + \epsilon_\lambda M_\lambda}{M_1 + M_2 + \dots + M_\lambda}$$

In der Tat : Wegen $\epsilon_1 \geq \epsilon_2 \geq \epsilon_3 \geq \dots$ ist

7) $$\frac{\epsilon_k M_k + \epsilon_{k+1} M_{k+1} + \dots + \epsilon_\lambda M_\lambda}{M_k + M_{k+1} + \dots + M_\lambda} \leq \epsilon_k \leq \epsilon_{k-1}$$

8) $$\epsilon_{k-1} = \frac{\epsilon_{k-1} M_1 + \dots + \epsilon_{k-1} M_{k-1}}{M_1 + \dots + M_{k-1}} < \frac{\epsilon_1 M_1 + \dots + \epsilon_{k-1} M_{k-1}}{M_1 + \dots + M_{k-1}}$$

Sind a, b, c und d beliebige positive Zahlen, so folgt aus $\frac{a}{b} < \frac{c}{d}$: $ad < bc$ oder $ab + ad < ab + bc$, also $\frac{a}{b} < \frac{a + c}{b + d}$. Setzt man den Bruch auf der linken Seite von (7) gleich $\frac{a}{b}$ und den letzten Bruch von (8) gleich $\frac{c}{d}$, so ergibt sich, dass die rechte Seite von (6) gleich $\frac{a + c}{b + d}$ ist. Damit ist (6) bewiesen.

Nun zum Beweis von (5) !

Aus (α) und Ziffer 2 folgt für $\lambda \geq k$, dass sich die Periode von $F^{(\lambda+1)}$ zyklisch an $F^{(\lambda)}$ anschliesst, so dass das zu $E^{[k]}$ gehörende Feld aus vollständigen reinen Perioden der zu den einzelnen $F^{(k+1)}$, $F^{(k+2)}$, ... gehörenden Feldern besteht. Es ist :

(9) $$H_n(E^{[k]}) = H_{M_k}(F^{(k)}) + H_{M_{k+1}}(F^{(k+1)}) + \dots + H_{M_\lambda}(F^{(\lambda)}) + H_z(F^{(\lambda+1)}) + H_{z'}^{(\lambda+1)} ,$$

wenn $\mu_\lambda - \mu_{k-1} \leq n < \mu_\lambda + 1 - \mu_{k-1}$

$$z = \left[\frac{n - \mu_\lambda + \mu_{k-1}}{P_{\lambda+1}}\right] P_{\lambda+1} \quad ; \quad z' = n - \mu_\lambda + \mu_{k-1} - z$$

gesetzt wird und $H_{z'}^{(\lambda+1)}$ die Häufigkeit der betrachteten Merkmalverbindung bei den letzten z' , aus Gliedern von

$F^{(\lambda+1)}$ gebildeten Spalten ist. Es ist

$$0 \leq z' < p_{\lambda+1} \quad \text{und} \quad 0 \leq H_{z'}^{(\lambda+1)} \leq z' < p_{\lambda+1} \ .$$

Aus (9) folgt

10)
$$\left| h_n(E^{[k]}) - \overline{w} \right| = \frac{M_k}{n} \left| \overline{w}^{(k)} - \overline{w} \right| + \frac{M_{k+1}}{n} \left| \overline{w}^{(k+1)} - \overline{w} \right| + \ldots + \frac{M_\lambda}{n} \left| \overline{w}^{(\lambda)} - \overline{w} \right| + \frac{z}{n} \left| \overline{w}^{(\lambda+1)} - \overline{w} \right| + \frac{z'}{n} \left| \frac{H_{z'}^{(\lambda+1)}}{z'} - \overline{w} \right| ,$$

da wegen (α) und den Eigenschaften der nach Ziffer 2 konstruierten Folgen $F^{(K)}$ für $K \geq k$

$$h_{M_K}(F^{(K)}) = \overline{w}^{(K)} \qquad \text{ist.}$$

Aus (3) und (10) ergibt sich, da $\frac{z}{n} < 1$ und $\epsilon_{\lambda+1} \leq \epsilon_\lambda$,

$$\left| h_n(E^{[k]}) - \overline{w} \right| < \frac{M_k}{n} \epsilon_k + \frac{M_{k+1}}{n} \epsilon_{k+1} + \ldots + \frac{M_\lambda}{n} \epsilon_\lambda + \epsilon_\lambda + \frac{p_{\lambda+1}}{\mu_\lambda} \frac{\mu_\lambda}{n}$$

$$< \frac{M_k \epsilon_k + \ldots + M_\lambda \epsilon_\lambda}{M_k + \ldots + M_\lambda} + \epsilon_\lambda + \epsilon_\lambda \frac{\mu_\lambda}{\mu_\lambda - \mu_{k-1}}$$

(wegen (γ) und $n > \mu_\lambda - \mu_{k-1}$)

$$< 2\epsilon_\lambda + \epsilon_\lambda + 2\epsilon_\lambda = 5\epsilon_\lambda \text{ , wegen (6), (β) und (δ) .}$$

Ist $\epsilon \geq 5\epsilon_\lambda$, so genügt es $N \geq \mu_\lambda$ zu wählen. Damit ist (5) bewiesen.

4. In entsprechender Weise lässt sich die Folge E zusammenstückeln, wenn für eine rationale Verteilung die Z-Folge konstruiert werden soll.

Schliesslich weisen wir noch darauf hin, dass für den Fall $w_1 + w_2 + w_3 + \ldots < 1$ ganz entsprechend wie beim Existenzbeweis für W-Folgen vorgegangen werden kann, da der Unempfindlichkeitssatz auch für Z-Folgen gilt.

§ 16. B - Folgen .

Der Vollständigkeit halber führen wir noch die Definition der Bernoullischen W-Folge an :

Def.VII. Eine E-Folge E soll in Bezug auf die endlich oder abzählbar unendlich vielen einander ausschliessenden Merkmale $m_1, m_2, m_3, \ldots$ eine Bernoullische W-Folge, abgekürzt B-Folge heissen, wenn folgende Bedingung für $k = 1, 2, 3, \ldots$ erfüllt ist : In dem abbrechenden Feld

$E_1 , E_{k+1} , E_{2k+1} , \ldots$

$E_2 , E_{k+2} , E_{2k+2} , \ldots$

$\ldots\ldots\ldots\ldots\ldots\ldots$

$E_k , E_{2k} , E_{3k} , \ldots$

existiert die Wahrscheinlichkeit dafür, dass in einer Spalte des Feldes in der x_1-ten Zeile m_{r_1}, in der x_2-ten Zeile m_{r_2}, ... , in der x_ν-ten Zeile m_{r_ν} vorkommt, und ist gleich $w_{r_1} w_{r_2} \ldots w_{r_\nu}$, unabhängig von $1 \leq x_1 < x_2 < \ldots$

$\ldots < x_\nu \leq k$. Dies soll für alle Merkmalkombinationen $m_{r_1}\, m_{r_2} \ldots m_{r_\nu}$ gelten. Man beachte den Unterschied des betrachteten Feldes gegenüber dem bei Z-Folgen.

Von den Eigenschaften der B-Folgen soll hier zunächst die Frage der Unempfindlichkeit der B-Folgen behandelt werden. In der allgemeinen Form des Unempfindlichkeitssatzes, wie er für E-Folgen, W-Folgen, X-Folgen und Z-Folgen gilt, wurde bisher der Unempfindlichkeitssatz für B-Folgen nicht bewiesen. Es gelten aber wenigstens die beiden folgenden Sätze, bei deren Beweis wir uns an die Bezeichnungen und Beweise bei Kamke und Copeland anlehnen :

atz 13. Ist E : E_1, E_2, E_3, ... eine B-Folge mit der Verteilung w_1, w_2, w_3, ... und kommen die Elemente E_{ν_1} , E_{ν_2} , E_{ν_3} , ... in E mit der Grenzhäufigkeit 0 vor, so ist auch die durch Abänderung der Elemente E_{ν_1} , E_{ν_2} , E_{ν_3} , ... entstehende Ersatzfolge eine B-Folge mit der Verteilung w_1, w_2, w_3, ...

Da sich durch das Ersetzen einer mit der Grenzhäufigkeit 0 vorkommenden Anzahl von Elementen E_ν für beliebiges k die Folge $(\overset{r}{k}E)$ ebenfalls in höchstens mit der Grenzhäufigkeit 0 vorkommenden Elementen ändert, so gilt dies nach dem allgemeinen Additionsgesetz[13] auch für eine beliebige Ver-

---

13) vgl. E.Kamke, a.a.O.,S.23,Satz 4.

bindungsfolge $\binom{r_1}{k}E) \binom{r_2}{k}E) \dots \binom{r_\nu}{k}E)$

mit $1 \leqq r_1 < r_2 < \dots < r_\nu \leqq k$. Es sind also die Voraussetzungen des Satzes 1 gegeben, sodass die Bedingung der Def.VII auch für die Ersatzfolge erfüllt ist.

Satz 14. Werden in einer B-Folge E endlich viele Elemente fortgelassen, hinzugefügt oder durch andere ersetzt, so ist auch die entstehende Folge E' eine B-Folge mit der gleichen Verteilung.

Der Beweis für den Fall zweier Merkmale ist bei Kamke[14] und Copeland[15] gegeben. Im allgemeinen Fall ist an Stelle der Formel (24) bei Kamke[14] die Beziehung

$$w\left\{\binom{g+r_1}{k}E) \binom{g+r_2}{k}E) \dots \binom{g+r_\nu}{k}E), m'm'' \dots m^{(\nu)}\right\} = w'w'' \dots w^{(\nu)}$$

zu beweisen. Der Beweis ist wörtlich gleich dem bei Kamke, wenn überall statt der Merkmalverbindung mm ... m (ν Merkmale) die Merkmalverbindung $m'm'' \dots m^{(\nu)}$ und statt der Wahrscheinlichkeit w^ν die Wahrscheinlichkeit $w'w'' \dots w^{(\nu)}$ gesetzt wird. Dabei kann n_0 und m_0 für festes k und für ein bestimmtes Indexsystem $1 \leqq r_1 < r_2 < \dots < r_\nu \leqq k$ so gewählt werden, dass die Ungleichung

---

14) vgl. E.Kamke, a.a.O.,S.140/142.

---

15) vgl. Copeland, American Journal of Mathematics 50 (1928) S.547.

$$\left| h_m\left\{ \binom{R_1}{k}E) \binom{R_2}{k}E) \dots \binom{R_\nu}{k}E) \, , \, m'm'' \dots m^{(\nu)} \right.\right.$$
$$\left.\left. - w'w'' \dots w^{(\nu)} \right\} \right| < \epsilon$$

für alle endlich vielen Kombinationen (k ist fest !) $m'm'' \dots m^{(\nu)}$ und für alle $m \geq m_0$ gilt.

An diesen Satz knüpfen wir noch eine Bemerkung : Bedeutet $E^{[\rho]}$ diejenige Folge, die aus E durch Streichung der ersten $\rho-1$ Glieder hervorgeht, so folgt aus Satz 14, dass mit E auch $E^{[\rho]}$ eine B-Folge mit gleicher Verteilung ist.

Wir werden noch zeigen, dass die B-Folgen ein Spezialfall der Z-Folgen sind :

atz 15. Jede B-Folge ist eine Z-Folge.

Für beliebiges k gilt nämlich, wenn $E^{(\rho)}$ wieder die in der ρ-ten Zeile des Z_k-Feldes stehende Folge bedeutet, für $\nu \leq k \, , \; 1 \leq r_1 < r_2 < \dots < r_\nu \leq k$:

$$E^{(r_1)}E^{(r_2)} \dots E^{(r_\nu)} = \binom{r_1}{k}E) \binom{r_2}{k}E) \dots \binom{r_\nu}{k}E)$$
$$+ \binom{r_1+1}{k}E) \binom{r_2+1}{k}E) \dots \binom{r_\nu+1}{k}E) + \dots$$
$$+ \binom{r_1+k-1}{k}E) \binom{r_2+k-1}{k}E) \dots \binom{r_\nu+k-1}{k}E)$$

Da E eine B-Folge sein soll, so sind die auf der rechten Seite stehenden Folgen W-Folgen mit der Wahrscheinlichkeit $w'w'' \dots w^{(\nu)}$ in Bezug auf eine beliebige Merkmalkombination $m'm'' \dots m^{(\nu)}$; dies folgt aus Satz 14. Auf der linken Seite

steht die Folge $E^{(r_1)}\ E^{(r_2)}\ \ldots\ E^{(r_\nu)}$, die eine Mischung der rechts stehenden Folgen ist und also ebenfalls die Wahrscheinlichkeit $w'w''\ \ldots\ w^{(\nu)}$ in Bezug auf $m'm''\ \ldots\ m^{(\nu)}$ hat.

Für B-Folgen gelten daher die Bernoullischen Sätze und zwar 1. in der hier betrachteten schwachen Form, wo man in der Folge alle Serien gleicher Länge beobachtet, ferner 2. in der stärkeren Form, wo man nur aufeinanderfolgende Serien gleicher Länge beobachtet[16] .

§ 17. Zusammenhänge .

1. Wir sind von der Definition der allgemeinen E-Folge ausgegangen und dadurch, dass wir die E-Folgen gewissen Forderungen unterwarfen, zu speziellen Folgen gelangt, den X-Folgen, W-Folgen, Z-Folgen und B-Folgen. Von den B-Folgen, die ein Spezialfall der Z-Folgen sind, sehen wir zunächst ab. Die Art der Spezialisierung können wir in eine Grenzwertforderung und in eine Forderung über die Auswahlfolgen einteilen, wobei diese Forderungen in Analogie zu dem Grenzwertaxiom und dem Regellosigkeitsaxiom bei v.Mises gesetzt werden können.

---

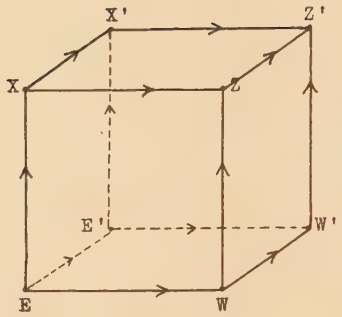
16) vgl. E.Kamke, a.a.O.,S.53.

Verlangt man von einer E-Folge, dass der Grenzwert der relativen Häufigkeit konvergiere, das W-Band also auf einen Punkt zusammenschrumpfe, so kommt man zu den W-Folgen. Verlangt man von den W-Folgen, dass sämtliche Auswahlfolgen unendliche Folgen sind und dass das kleinste Intervall, in dem die W-Bänder der Folge und sämtlicher Auswahlfolgen enthalten sind, aus einem Punkt bestehe, so erhält man die Z-Folge.

Setzt man von einer E-Folge voraus, dass alle Auswahlfolgen unendliche Folgen sind, so ist die Folge X-Folge. Verlangt man, dass das kleinste Intervall, in dem die W-Bänder der X-Folge und aller Auswahlfolgen enthalten sind, nur einen Punkt enthält, so wird aus der X-Folge die Z-Folge.

Jede der betrachteten Folgen kann man nun dadurch spezialisieren, dass man verlangt, dass die W-Bänder der Folge und etwa betrachteter Auswahlfolgen in einem ganz im Innern von $[0,1]$ liegenden Intervall enthalten sein sollen. Diese Spezialisierung deuten wir durch einen Strich an. Z' ist also eine Z-Folge, deren Wahrscheinlichkeiten sämtlich grösser als 0 und kleiner als 1 sind.

Eine schematische Ordnung der Folgen erhalten wir, wenn wir die Folgenarten in der nachstehenden Weise andeuten :



Die Pfeile deuten die Richtung der Spezialisierung an, und zwar
Pfeile parallel zu EW eine Grenzwertforderung,
Pfeile parallel zu EX eine Forderung über Auswahlfolgen,
Pfeile parallel zu EE'eine Forderung über den Ausschluss trivialer Folgen .

2. Wir haben gezeigt, dass - mit Ausnahme der B-Folgen - für alle Folgen der Unempfindlichkeitssatz gilt.

Die 1. Bernoullische Aussage gilt für die X-Folgen und die Z-Folgen; dagegen konnte die 2. Bernoullische Aussage für X'-Folgen nicht bewiesen werden; andererseits braucht man nicht die vollen Voraussetzungen über die Z'-Folgen zum Beweis der 2. Bernoullischen Aussage, da sich deren Richtigkeit bereits für die Poisson'schen X'-Folgen ergeben hat.

3. Die B-Folgen sind ein Spezialfall der Z'-Folgen. Für die B-Folgen gilt die 1. und 2. Bernoullische Aussage, sogar in der stärkeren Form als bei Z'-Folgen. Für die Behandlung der Aufgaben der Wahrscheinlichkeitsrechnung wurden bisher im wesentlichen B-Folgen zugrunde gelegt. Tatsächlich gibt es - z.B. bei der Behandlung von Aufgaben über Glücksspiele - eine Reihe von Aufgaben, zu deren Lösung die bekannten Eigenschaften der Z-Folgen nicht ausreichen. Andererseits sind bei Anwendung der Z-Folgen manche Aufgaben, wie z.B. die Behandlung der Iterationen einfacher zu lösen. Soweit die Z-Folgen für die Anwendungsgebiete der Wahrscheinlichkeitsrechnung ausreichen, sind sie den B-Folgen jedenfalls solange vorzuziehen, als der allgemeine Unempfindlichkeitssatz für B-Folgen nicht bewiesen ist; dies auch solange, als man nicht weiss, ob die B-Folgen gegen Vorgängerauswahl unempfindlich sind, d.h. ob die Auswahlfolgen einer B-Folge wieder B-Folgen sind, was hier für Z-Folgen bewiesen wurde.

   Darüber hinaus wird man da, wo es nur auf die 1. Bernoullische Aussage ankommt, wie dies z.B. bei der Anwendung auf die Statistik der Fall ist, sogar auf die X-Folgen zurückgreifen können, da man hier die nicht nachprüfbare Forderung der Konvergenz vermeidet.

4. Die Menge aller B-Folgen mit vorgeschriebener Verteilung hat die grösstmögliche Mächtigkeit, nämlich die Mächtig-

keit des Kontinuums[17]. Da die B-Folgen einen Spezialfall aller übrigen hier betrachteten Folgenarten darstellen, so haben auch die Mengen aller X-Folgen, W-Folgen und Z-Folgen die gleiche Mächtigkeit. Uebrigens würde es grundsätzlich genügen, die Existenz der B-Folgen nachzuweisen, da damit die Existenz aller übrigen Folgen gesichert ist; jedoch ist der Existenzbeweis für Z-Folgen einfacherer Art als der für B-Folgen.

Dass die Menge der W-Folgen und der Z-Folgen nicht identisch ist, zeigt das Beispiel einer W-Folge, die nicht Z-Folge ist. In der Folge 0, 1, 0, 1, 0, 1, ... ist die Wahrscheinlichkeit für das Auftreten der 0 vorhanden und gleich 1/2. Wählt man die Glieder mit dem Vorgänger 1 aus, so erhält man eine nur aus Elementen 0 bestehende Folge, die zwar wieder W-Folge ist, aber die eine von 1/2 verschiedene Wahrscheinlichkeit, nämlich 1 hat.

Die Frage, ob es Z-Folgen gibt, die nicht B-Folgen sind, muss hier unbeantwortet bleiben.

---

17) vgl. E.Kamke, a.a.O.,S.139 f.

Literaturverzeichnis.

A.H.Copeland — American Journal of Mathematics 50 (1928) S.535/552.

E. Czuber — Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung. 2 Bde. 4. Auflage. Leipzig und Berlin 1924 und 1928.

E. Kamke — Einführung in die Wahrscheinlichkeitstheorie. Leipzig 1932.

R.v.Mises — Wahrscheinlichkeit, Statistik und Wahrheit. Wien 1928.

R.v.Mises — Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig und Wien 1931.

K.Popper — Logik der Forschung. Wien 1935.

H.Reichenbach — Wahrscheinlichkeitslehre. Leiden 1935.

E.Tornier — Journal für die reine und angewandte Mathematik 160 (1929) S.177/198 und 163 (1930) S.45/64.

Inhaltsverzeichnis.

	Seite
Einleitung	1
I. X - Folgen	
§ 1. E-Folgen	3
§ 2. Das Rechnen mit W-Bändern	5
§ 3. Unempfindlichkeitssatz	10
§ 4. Additionssätze	13
§ 5. Der V_k-Bereich	16
§ 6. X-Folgen	22
§ 7. Die 1. Bernoullische Aussage für Verteilungssysteme	32
§ 8. Ueber die 2. Bernoullische Aussage für ein Verteilungssystem	41
§ 9. Die Bernoullischen Aussagen für X-Folgen ...	63
§ 10. Unempfindlichkeit und Existenz von X-Folgen	68
II. Z - Folgen	
§ 11. W-Folgen	71
§ 12. Z-Folgen	81
§ 13. Sätze über Z-Folgen	86
§ 14. Satz über die Auswahlfolge einer Z-Folge ...	89
§ 15. Der Existenzsatz für Z-Folgen	100
§ 16. B-Folgen	110
§ 17. Zusammenhänge	114
Literaturverzeichnis	119
Inhaltsverzeichnis	120

Lebenslauf

Am 24. April 1906 wurde ich, Gerhard Schauffler, als Sohn des Pfarrers Johannes Schauffler in Stuttgart geboren. Nach dem Besuch der Volksschule in Berghülen, der Höheren Knabenschule in Korntal und des Gymnasiums in Ulm bestand ich im Frühjahr 1923 die Reifeprüfung.

Anschließend studierte ich an den Universitäten Tübingen, München, Göttingen und Frankfurt a. M. Mathematik und Versicherungswissenschaft. Im Herbst 1931 bestand ich die Abschlußprüfung am Seminar für Versicherungswissenschaft in Frankfurt a. M. Seitdem bin ich als Versicherungsmathematiker in der Praxis tätig.

Meine akademischen Lehrer waren vor allem die Herren Professoren Carathéodory, Hilpert, Knopp, Kommerell, Perron, Schönhardt und Tietze. Ihnen allen bin ich zu herzlichem Dank verpflichtet; ganz besonderen Dank schulde ich Herrn Professor Dr. Aumann für die wertvollen Anregungen und die rege Anteilnahme bei der Ausarbeitung dieser Untersuchung.